ELSEVIER

Journal of Econometrics 108 (2002) 63–99

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Testing for stationarity with a break

Eiji Kurozumi*

Graduate School of Economics, Hitotsubashi University, Naka 2-1, Kunitachi, Tokyo, 186-8601, Japan

Received 15 May 2000; revised 26 June 2001; accepted 13 July 2001

Abstract

In this paper, we investigate a test for the null hypothesis of trend stationarity with a structural change against a unit root. We derive the limiting distribution of an Lagrange Multiplier (LM) test statistic and its characteristic function under a sequence of local alternatives. The local limiting power of the LM test depends on the persistence of the stationary component of the process, and the more persistent the process, the less powerful is the test statistic. We also propose a test statistic that does not depend on the fraction of the pre-break points to the sample size under the null hypothesis, which we call the PS test. Though it is convenient for the critical point not to depend on the break point, the PS test is found to be less powerful than the LM test under the alternative close to the null hypothesis. Finite sample simulations show that when the break point is known, the LM test tends to be oversized when the process is rather persistent, while the size distortion of the PS test is not so pronounced. On the other hand, the empirical sizes of both tests are close to the nominal one when the break point is estimated by the least-squares method, though the power decreases compared with the known break point case.

JEL classification: C11; C22

Keywords: Hypothesis testing; Structural change; Trend stationary

1. Introduction

Economists have been concerned to know whether there is persistence of a unit root in a time series, and testing for a unit root has an important role

*Corresponding author. Tel.: +81-42-580-8485; fax: +81-42-580-8195.
E-mail address: kurozumi@stat.hit-u.ac.jp (E. Kurozumi).

0304-4076/02/$ - see front matter
PII: S0304-4076(01)00106-3

in practical analyses. From many empirical studies following the work of Nelson and Plosser (1982), it has been concluded that many macroeconomic time series have a unit root, and more or less persistence was found in such data.

However, Perron (1989) showed that the Dickey–Fuller (D–F) test cannot reject the null hypothesis of a unit root under the alternative of trend stationarity with a structural change. That is, the D–F test does not have nontrivial power against stationarity with a break. Perron (1989, 1990) and Perron and Vogelsang (1992b) developed testing procedures against such a process with a known break point, whereas, with an unknown break point, unit root tests have been proposed in the works of Banerjee et al. (1992), Perron (1997), Perron and Vogelsang (1992a), Vogelsang and Perron (1998) and Zivot and Andrews (1992), among others.

All the above papers consider the null hypothesis of a unit root against the alternative of trend stationarity with a break, but in this paper, we consider the testing problem in the reverse direction. That is, we consider the null of trend stationarity with a break against the alternative of a unit root. Considering such a hypothesis may be important for researchers who have interest in a unit root. In view of the mechanism by which classical hypothesis testing is carried out, acceptance of the null of a unit root does not necessarily imply the existence of a unit root, but the unit root hypothesis would be supported more strongly if the null of stationarity is rejected by the test constructed in the reverse direction. In this sense, we may see that the test for the null of stationarity complements the unit root test. On the other hand, when the null of a unit root is rejected, we cannot conclude that the process is trend stationary, as the unit root tests considered in the above papers may have power not only against a stationary process with a break but also against more general alternatives. Then, once the null of a unit root is rejected, the test for the null of trend stationarity with a break becomes of primary interest.

We propose such a testing procedure with a Lagrange Multiplier (LM) test statistic for four patterns of the break shift. Tests for stationarity when no break have been proposed in the literature by, for example, Kwiatkowski et al. (1992b) (hereafter KPSS) and Leybourne and McCabe (1994). The main difference between these two tests is that the former uses a nonparametric correction to exclude the nuisance parameter from the limiting distribution, while the latter restricts the model to the finite order ARMA model. Because we use a nonparametric method to exclude the nuisance parameter, the LM test in this paper may be seen as an extension of the KPSS test to the trend break case. We suppose that the fraction of the pre-break points to the sample size is constant, and, as in the tests for the null of a unit root with a possible break, the limiting distribution of our test depends on this fraction. We also propose a test statistic using the same method as Park and Sung (1994) that does not depend on the fraction under the null hypothesis (we call this the PS test). The limiting properties of the tests proposed in this paper are compared under a sequence of local alternatives, and, as suggested by theory that the LM test is locally best invariant (LBI) under the assumption of normality, the limiting power of the LM test is found to dominate that of the PS test under the alternative close to the null, although this is not always the case when the local alternatives diverge from the null.

The plan of the paper is as follows. Section 2 sets up the model and the assumptions. Two test statistics are proposed in Section 3 and their limiting properties are investigated when the break date is known. The unknown break point case is treated in Section 4. Finite sample properties are investigated in Section 5 and the tests proposed in the paper are applied to US macroeconomic data in Section 6. Section 7 concludes the paper.

2. The model and the testing problem

Let us consider the following error-components model:

$$y_t = z_t'\beta + x_t, \qquad x_t = \gamma_t + u_t, \qquad \gamma_t = \gamma_{t-1} + \varepsilon_t, \qquad u_t = \sum_{j=0}^{\infty} \alpha_j v_{t-j}, \tag{1}$$

where z_t denotes a deterministic component that includes a trend break, $\{v_t, \varepsilon_t\}'$ are jointly independently and identically distributed with $E[v_t^2] = \sigma_v^2 > 0$ and $E[\varepsilon_t^2] = \sigma_\varepsilon^2 \geqslant 0$, and $\{u_t\}$ and $\{\varepsilon_t\}$ are independent. We assume that $\{v_t, \varepsilon_t\}'$ has finite $2+\delta$th moment for $\delta > 0$. We also assume that $\sum_{j=1}^{\infty} j|\alpha_j| < \infty$ and $\alpha \equiv \sum_{j=0}^{\infty} \alpha_j \neq 0$. We set $t = 1, \ldots, T$ and $\gamma_0 = 0$ without loss of generality as z_t includes a constant term as defined below. We suppose that a structural change has occurred at time T_B $(1 < T_B < T)$, and that $\omega = T_B/T$ is fixed. For the deterministic component, z_t, we consider the following four cases:

Case 0: a constant with a break; $z_t = [1, DU_t]'$,

Case 1: a constant with a break and a linear trend; $z_t = [1, DU_t, t/T]'$,

Case 2: a constant with no break and a linear trend with a break; $z_t = [1, t/T, DT_t]'$,

Case 3: a constant and a linear trend, both with a break; $z_t = [1, DU_t, t/T, DT_t]'$,

where $DU_t = 1(t > T_B)$ and $DT_t = 1(t > T_B) \times (t - T_B)/T$ with $1(\cdot)$ denoting an indicator function. Case 0 corresponds to the model without a linear trend such as an interest rate and the purchasing power parity as discussed in Perron (1990) and Perron and Vogelsang (1992a, b), whereas Cases 1–3 apply to a model with a linear trend such as gross domestic product and many other macroeconomic variables. Perron (1989) called Case 1 the "crash model" and

Case 2 the "changing growth model". Case 3 allows for a "sudden change in level followed by a different growth path".

Basically, we will investigate the above "additive outlier model" with a known break point. That is, we suppose that we know when the structural change occurred and that a shock affects the observations only at one time, but we will later discuss the case of the "innovational outlier model", in which the structural change disturbs the variables with lagged effects. The unknown break point case is discussed in Section 4.

Model (1) can be expressed as a vectorized model by stacking each variable,

$$y = Z\beta + x, \qquad x = \gamma + u, \qquad \gamma = L\varepsilon,$$

where, e.g., $y = [y_1, \ldots, y_T]'$ and L is a lower triangular matrix with lower elements 1's,

$$L = \begin{bmatrix} 1 & & \mathbf{0} \\ \vdots & \ddots & \\ 1 & \cdots & 1 \end{bmatrix}.$$

To test for the null hypothesis of stationarity with a break, we consider the following testing problem:

$$\mathrm{H}_0\colon \rho = \frac{\sigma_\varepsilon^2}{\sigma_v^2} = 0 \quad \text{vs.} \quad \mathrm{H}_1\colon \rho = \frac{c^2}{T^2}, \tag{2}$$

where c is a constant. Then, under the null hypothesis, $\sigma_\varepsilon^2 = 0$ so that $\{y_t\}$ is trend stationary with a break. On the other hand, under H_1, $\{x_t\}$ contains a unit root component $\{\gamma_t\}$ so that $\{y_t\}$ becomes a unit root process with a break. Note that, in our model (1), the order T^{-2} is required for the possible local asymptotic analysis under a sequence of local alternatives. This is because the test statistic can be expressed as a function of the partial sums of $\tilde{x}_t$ as shown in (3), that is, $\sum_{t=1}^{j} \tilde{x}_t = \sum_{t=1}^{j} \tilde{u}_t + \sum_{t=1}^{j} \tilde{\gamma}_t$ where $\tilde{u}_t$ and $\tilde{\gamma}_t$ are regression residuals of u_t and γ_t on z_t. The first term is of order $T^{1/2}$ as $\{u_t\}$ is second order stationary whereas the latter is of order $T^{3/2}$ if we consider the fixed alternative, so that the latter dominates the former and the test statistic diverges to infinity. To proceed with the local asymptotic analysis, both partial sums of $\tilde{u}_t$ and $\tilde{\gamma}_t$ have to be of the same order, $T^{1/2}$, and to establish this, we assume that ρ is of order T^{-2}. Notice that $\sum_{t=1}^{j} \tilde{\gamma}_t$ is $\mathrm{O_p}(T^{1/2})$ when $\rho = c^2/T^2$. By considering a sequence of local alternatives, not just one fixed alternative, we can derive local limiting power functions and investigate the properties of the test statistics by drawing such functions.

3. Testing for stationarity with a known break point

3.1. The LM test

For the testing problem (2), it is well known that when $u_t = v_t$, the LM test statistic is proportional to $\hat{\sigma}^{-2} y'MLL'My$ where $M = I_T - Z(Z'Z)^{-1}Z'$ and $\hat{\sigma}^2 = T^{-1}y'My$. See, for example, Kwiatkowski et al. (1992a) for the derivation. Note that under the assumption of normality, the LM test is equivalent to the LBI test as discussed in King and Hillier (1985).

Since the limiting distribution of the above LM test statistic depends on a nuisance parameter when u_t obeys a general process as in (1), we consider the following statistic:

$$S_T = \frac{1}{\tilde{\sigma}^2 T^2} y'MLL'My = \frac{1}{\tilde{\sigma}^2 T^2} \sum_{j=1}^{T} \left(\sum_{t=1}^{j} \tilde{x}_t \right)^2, \tag{3}$$

where

$$\tilde{\sigma}^2 = \sum_{i=-\ell}^{\ell} w(i,\ell) \frac{1}{T} \sum_{t=1}^{T-i} \tilde{x}_t \tilde{x}_{t+i} \tag{4}$$

with $\tilde{x}_t$ regression residuals of y_t on z_t,

$$\tilde{x}_t = y_t - z_t' \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \sum_{t=1}^{T} z_t y_t.$$

$w(i,\ell)$ is any kernel function that produces a nonnegative estimate of the long-run variance of $\{u_t\}$ and we assume $\ell = o_p(T^{1/2})$. The second expression of (3) is convenient for the practical calculation of the test statistic, though we use mainly the first expression for the theoretical explanation.

The following theorem provides the limiting distribution of S_T and its characteristic function for each case. For notational convenience, we define the following functional of a standard Brownian motion in generic form,

$$G(B;c^2) = \int_0^1 B(r)^2\,\mathrm{d}r - X(B)'\Lambda^{-1}X(B) + c^2 \int_0^1 \left(\int_0^r B(s)\,\mathrm{d}s - Z(r)'\Lambda^{-1}X(B) \right)^2 \mathrm{d}r,$$

where $B(\cdot)$ is a standard Brownian motion and $X(B)$ denotes a functional of $B(\cdot)$. As the null hypothesis is a special case of the alternative ($c = 0$), we give the result only under the alternative.

Theorem 1. Consider the model (1). (i) *For Cases* 0 *and* 3, *under a sequence of local alternatives*, H_1,

$$S_T \xrightarrow{d} \omega^2 G(B_1; c^2\omega^2/\alpha^2) + (1-\omega)^2 G(B_2, c^2(1-\omega^2)/\alpha^2) \tag{5}$$

and its characteristic function is expressed as

$$\begin{aligned}\phi(\theta; c) = {} & [D(\mathrm{i}\omega^2\theta + \sqrt{-\omega^4\theta^2 + 2\mathrm{i}c^2\omega^4\theta/\alpha^2}) \\ & \times D(\mathrm{i}\omega^2\theta - \sqrt{-\omega^4\theta^2 + 2\mathrm{i}c^2\omega^4\theta/\alpha^2})]^{-1/2} \\ & \times [D(\mathrm{i}(1-\omega)^2\theta + \sqrt{-(1-\omega)^4\theta^2 + 2\mathrm{i}c^2(1-\omega)^4\theta/\alpha^2}) \\ & \times D(\mathrm{i}(1-\omega)^2\theta - \sqrt{-(1-\omega)^4\theta^2 + 2\mathrm{i}c^2(1-\omega)^4\theta/\alpha^2})]^{-1/2},\end{aligned} \tag{6}$$

where $B_1(\cdot)$ *and* $B_2(\cdot)$ *are independent Brownian motions*, $\mathrm{i} = \sqrt{-1}$, *and*

(ia) *for Case* 0,

$$X(B) = \int_0^1 B(r)\,\mathrm{d}r, \quad Z(r) = r, \quad \Lambda = 1, \quad D(\lambda) = \frac{\sin\sqrt{\lambda}}{\sqrt{\lambda}},$$

(ib) *for Case* 3,

$$X(B) = \begin{bmatrix} \int_0^1 B(r)\,\mathrm{d}r \\ \int_0^1 rB(r)\,\mathrm{d}r \end{bmatrix}, \quad Z(r) = \begin{bmatrix} r \\ \frac{r^2}{2} \end{bmatrix}, \quad \Lambda = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{bmatrix},$$

$$D(\lambda) = \frac{12}{\lambda^2}(2 - \sqrt{\lambda}\sin\sqrt{\lambda} - 2\cos\sqrt{\lambda}).$$

(ii) *For Cases* 1 *and* 2, *under a sequence of local alternatives*, H_1,

$$S_T \xrightarrow{d} G(B; c^2/\alpha^2) \tag{7}$$

and its characteristic function is expressed as

$$\phi(\theta; c) = [D(\mathrm{i}\theta + \sqrt{-\theta^2 + 2\mathrm{i}c^2\theta/\alpha^2})D(\mathrm{i}\theta - \sqrt{-\theta^2 + 2\mathrm{i}c^2\theta/\alpha^2})]^{-1/2}, \tag{8}$$

where $B(\cdot)$ *is a standard Brownian motion and*

(iia) *for Case* 1,

$$X(B) = \begin{bmatrix} \int_0^1 B(r)\,\mathrm{d}r \\ \int_\omega^1 B(r)\,\mathrm{d}r \\ \int_0^1 rB(r)\,\mathrm{d}r \end{bmatrix}, \quad Z(r) = \begin{bmatrix} r \\ \mathrm{d}t_r \\ \frac{r^2}{2} \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} 1 & 1-\omega & \frac{1}{2} \\ 1-\omega & 1-\omega & \frac{1-\omega^2}{2} \\ \frac{1}{2} & \frac{1-\omega^2}{2} & \frac{1}{3} \end{bmatrix},$$

$$D(\lambda) = -12\frac{\sqrt{\lambda}\sin\sqrt{\lambda\omega^2}\sin\sqrt{\lambda(1-\omega)^2}+2(\sin\sqrt{\lambda}-\sin\sqrt{\lambda\omega^2}-\sin\sqrt{\lambda(1-\omega)^2})}{\lambda^{5/2}\omega(1-\omega)\{1-3\omega(1-\omega)\}},$$

(iib) *for Case* 2,

$$X(B) = \begin{bmatrix} \int_0^1 B(r)\,\mathrm{d}r \\ \int_0^1 rB(r)\,\mathrm{d}r \\ \int_\omega^1 (r-\omega)B(r)\,\mathrm{d}r \end{bmatrix}, \quad Z(r) = \begin{bmatrix} r \\ \frac{r^2}{2} \\ 1(r>\omega)\frac{(r-\omega)^2}{2} \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} 1 & \frac{1}{2} & \frac{(1-\omega)^2}{2} \\ \frac{1}{2} & \frac{1}{3} & \frac{(1-\omega)^2(\omega+2)}{6} \\ \frac{(1-\omega)^2}{2} & \frac{(1-\omega)^2(\omega+2)}{6} & \frac{(1-\omega)^3}{3} \end{bmatrix},$$

$$D(\lambda) = \frac{D_1(\lambda)+D_2(\lambda)+D_3(\lambda)}{\lambda^{7/2}\omega^3(1-\omega)^3},$$

with

$$D_1(\lambda) = \lambda\omega(1-\omega)\sin\sqrt{\lambda},$$

$$D_2(\lambda) = 2\{\sin\sqrt{\lambda\omega^2}+\sin\sqrt{\lambda(1-\omega)^2}-\sin\sqrt{\lambda} - \lambda^{1/2}(\omega\cos\sqrt{\lambda\omega^2}+(1-\omega)\cos\sqrt{\lambda(1-\omega)^2})\},$$

$$D_3(\lambda) = \lambda^{1/2}(\cos\sqrt{\lambda}+\cos\sqrt{\lambda\omega^2}\cos\sqrt{\lambda(1-\omega)^2}),$$

where $\mathrm{d}t_r = 1(r>\omega)\times(r-\omega)$.

Remark 1. The test statistic S_T is consistent in the sense that $\lim_{T\to\infty} P(S_T > x_*) \to 1$ as $c \to \infty$ where x_* is a critical point, because $G(B;c^2) \to \infty$ as $c \to \infty$. We can also show that S_T diverges to infinity when we consider the fixed alternative, $\rho > 0$, so that the LM test statistic S_T is consistent in the usual sense. This can be proved exactly in the same way as KPSS (1992b), that is, when $\rho > 0$ is fixed, $\{\gamma_t\}$ dominates $\{u_t\}$ and the order of $\sum_{t=1}^{j} \tilde{x}_t$

is $T^{3/2}$. Then, $\sum_j (\sum_t \tilde{x}_t)^2$ in Eq. (3) is of order T^4 while $\tilde{\sigma}^2$ is of order ℓT as shown in KPSS (1992b). As a result, the order of the test statistic (3) becomes T/ℓ and then the test statistic is consistent in the usual sense as $\ell = \mathrm{o_p}(T^{1/2})$.

Remark 2. From (5) and (7), we can see that the larger value of $|\alpha|$ entails the lower power. For example, when $\{u_t\}$ obeys an AR(1) process such as $u_t = au_{t-1} + v_t$, α^2 is $1/(1-a)^2$, so that the larger the value of a becomes, or the more persistent the process becomes, the more slowly does the power function increase as a function of c.

Remark 3. For Cases 0 and 3, the limiting distribution is expressed as the sum of two independent functionals, $G(B_1)$ and $G(B_2)$, so that its characteristic function is expressed as the product of two characteristic functions. This is because the test statistic S_T can be expressed as the sum of two functions, one being a function depending on the observation before the break and the other being so after the break. The proof is in the Appendix. As S_T for Cases 1 and 2 cannot be expressed in such a form, its characteristic function becomes a little complicated. Though the limiting distributions for Cases 0 and 3 can also be expressed as (7), expression (5) may be more useful for understanding why their characteristic functions have the form as (6).

Remark 4. Under the null hypothesis, $c = 0$ so that S_T converges in distribution to

$$\omega^2 \left(\int_0^1 B_1(r)^2 \, \mathrm{d}r - X(B_1)' \Lambda^{-1} X(B_1) \right) + (1-\omega)^2 \left(\int_0^1 B_2(r)^2 \, \mathrm{d}r - X(B_2)' \Lambda^{-1} X(B_2) \right), \tag{9}$$

for the Cases 0 and 3, and

$$\int_0^1 B(r)^2 \, \mathrm{d}r - X(B)' \Lambda^{-1} X(B), \tag{10}$$

for Cases 1 and 2. Their characteristic functions can be expressed more compactly as

$$\phi(\theta, 0) = [D(2\mathrm{i}\omega^2\theta) D(2\mathrm{i}(1-\omega)^2\theta)]^{-1/2}, \tag{11}$$

for Cases 0 and 3, and

$$\phi(\theta, 0) = [D(2\mathrm{i}\theta)]^{-1/2}, \tag{12}$$

for Cases 1 and 2.

Table 1
Percent points of the null distribution of the LM test

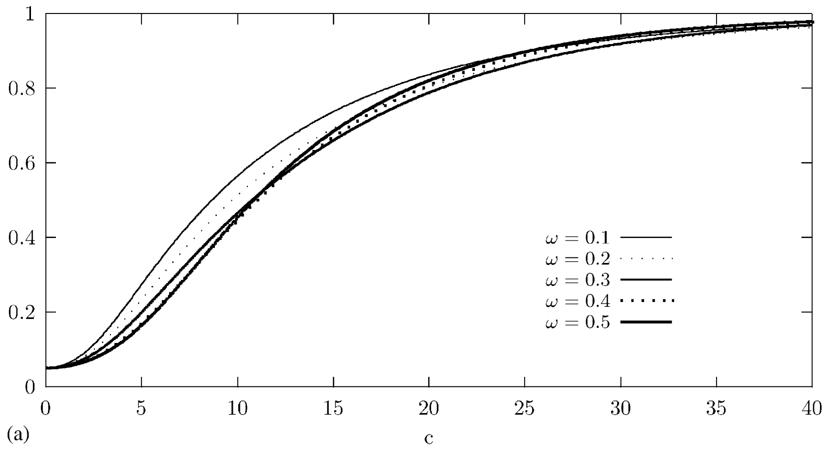
	0.01	0.05	0.1	0.5	0.9	0.95	0.99
a. Case 0							
$\omega = 0.1$	0.02160	0.03123	0.03892	0.09797	0.28299	0.37538	0.60388
$\omega = 0.2$	0.02049	0.02895	0.03548	0.08302	0.22915	0.30212	0.48265
$\omega = 0.3$	0.02001	0.02796	0.03396	0.07440	0.18678	0.24247	0.38052
$\omega = 0.4$	0.01978	0.02749	0.03326	0.07050	0.16007	0.20106	0.30162
$\omega = 0.5$	0.01971	0.02736	0.03305	0.06939	0.15176	0.18688	0.26842
b. Case 1							
$\omega = 0.1$	0.01544	0.02057	0.02426	0.04680	0.09840	0.12162	0.17821
$\omega = 0.2$	0.01517	0.02005	0.02350	0.04343	0.08537	0.10376	0.14839
$\omega = 0.3$	0.01525	0.02019	0.02370	0.04412	0.08579	0.10304	0.14291
$\omega = 0.4$	0.01541	0.02050	0.02415	0.04623	0.09736	0.12080	0.17842
$\omega = 0.5$	0.01549	0.02066	0.02439	0.04741	0.10551	0.13378	0.20405
c. Case 2							
$\omega = 0.1$	0.01536	0.02064	0.02448	0.04816	0.10263	0.12716	0.18696
$\omega = 0.2$	0.01441	0.01907	0.02242	0.04267	0.08879	0.10956	0.16020
$\omega = 0.3$	0.01394	0.01825	0.02129	0.03907	0.07815	0.09563	0.13829
$\omega = 0.4$	0.01371	0.01784	0.02073	0.03712	0.07138	0.08643	0.12299
$\omega = 0.5$	0.01364	0.01772	0.02056	0.03651	0.06909	0.08318	0.11727
d. Case 3							
$\omega = 0.1$	0.01463	0.01962	0.02325	0.04566	0.09724	0.12046	0.17704
$\omega = 0.2$	0.01331	0.01744	0.02039	0.03826	0.07903	0.09737	0.14208
$\omega = 0.3$	0.01267	0.01634	0.01889	0.03343	0.06485	0.07889	0.11308
$\omega = 0.4$	0.01237	0.01582	0.01817	0.03095	0.05570	0.06615	0.09122
$\omega = 0.5$	0.01228	0.01566	0.01796	0.03022	0.05267	0.06163	0.08216

From the above theorem, we can obtain the distribution function $F(x)$ in each case by inverting the characteristic function. Since the limiting distribution is nonnegative, we can calculate the percent points by numerical integration, using Lévy's inversion formula,

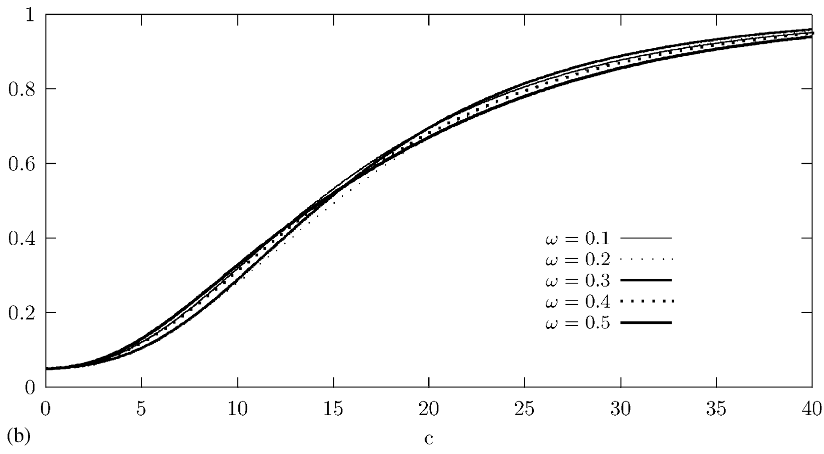
$$F(x) = \frac{1}{\pi}\int_0^{\infty} \mathrm{Re}\left[\frac{1-\mathrm{e}^{-\mathrm{i}\theta x}}{\mathrm{i}\theta}\phi(\theta; c)\right] \mathrm{d}\theta. \tag{13}$$

Especially for the null distribution, we set $c = 0$; that is, we use the characteristic function (11) or (12).

Table 1 reports the percent points for Cases 0–3. Because, as we can see from the characteristic function, the limiting distribution when $\omega = \omega^0$ is the same as when $\omega = 1 - \omega^0$, that is, it is symmetric around $\omega = 0.5$, we tabulate percentiles only for $\omega = 0.1$, 0.2, 0.3, 0.4 and 0.5. For $\omega > 0.5$, we can refer to the tables corresponding to $1 - \omega$.



(a)



(b)

Fig. 1. The limiting powers: (a) (Case 0); (b) (Case 1); (c) (Case 2); (d) (Case 3).

As in the case of the null distribution, the location of the break point, ω, also affects the limiting power properties. The limiting power function can also be calculated by numerical integration and is given by $1-F(x)$ as a function of c. Figs. 1a–d display the power functions for $\omega = 0.1$, 0.2, 0.3, 0.4 and 0.5 with $\alpha = 1$. For Cases 0, 2 and 3, the power for the smaller ω dominates that for the larger ω ($\leqslant 0.5$) near the null hypothesis. On the other hand, for Case 1, the power function for $\omega = 0.1$ is located above that for $\omega = 0.3$, but the case of $\omega = 0.5$ is most powerful among five values of ω when c is close to 0. These properties seem to be only for the small values of c, and as c increases, the above relation does not hold.

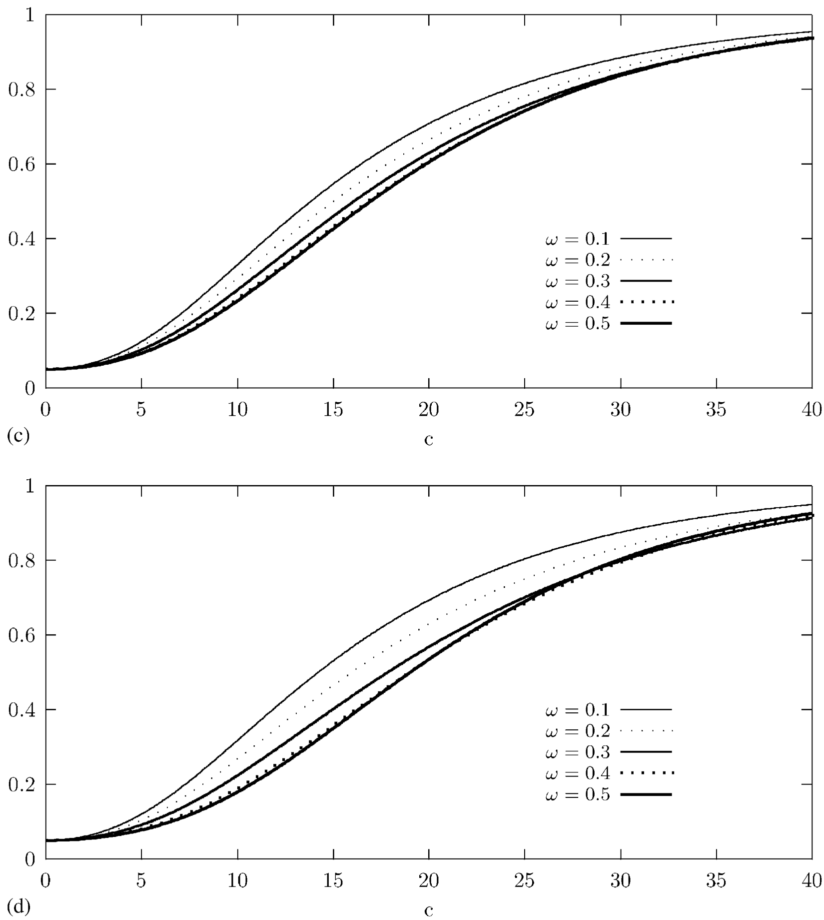


Fig. 1. (*continued*).

Next, we compare the limiting power functions of four cases for the fixed ω. Fig. 2 shows them for $\omega = 0.3$. As in many other tests, such as the D–F test, the more complicated the deterministic term becomes, the less powerful is the test statistic. We can see that the power function of Case 0 dominates the other three cases, and the test in Case 3 is the least powerful. These differences among power functions tend to diminish as the value of ω decreases to 0.1, and especially when $\omega = 0.1$, the power functions of Cases 1–3 are almost the same, though the power of Case 0 still dominates the others.

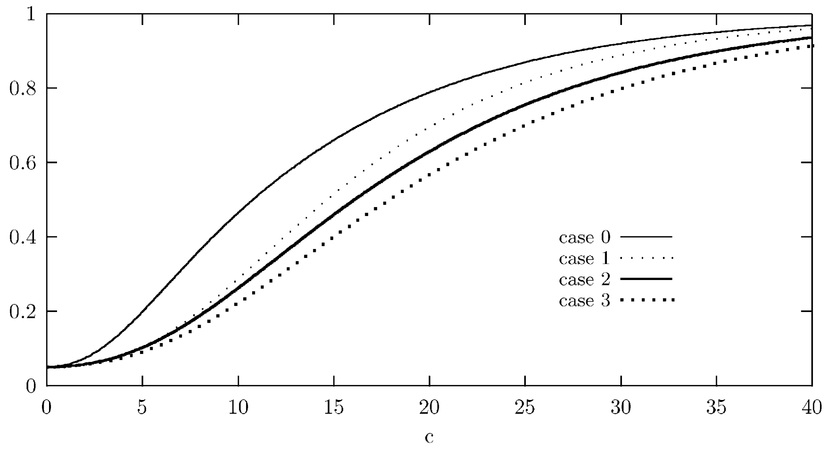


Fig. 2. The limiting powers ($\omega = 0.3$).

3.2. The test independent of the break point

As we can see from Theorem 1, the limiting distribution of the LM test statistic depends on the fraction of the pre-break points, and we tabulated the percent points of the null distribution for several ω's. In this section we consider the test statistic whose limiting distribution does not depend on the value of ω under the null hypothesis. As we can construct such a test only for Cases 0 and 3, that is, for the cases when there is one time break in all deterministic components included in the model, we consider only Cases 0 and 3 in this section. Fundamentally, our method is the same as that of Park and Sung (1994).

For Cases 0 and 3, the test statistic S_T is decomposed into the sum of two functions that are asymptotically independent each other, as discussed in Remark 3. As these functions consist of observations before and after the break date, respectively, the appropriate normalizing orders for these functions are not T^2 but T_B^2 and $(T - T_B^2)$. To accomplish these orders, we construct the weighted variable y_t^* following the idea of Park and Sung (1994).

$$y_t^* = \begin{cases} (T/T_B)y_t: & t = 1, \ldots, T_B, \\ [T/(T - T_B)]y_t: & t = T_B + 1, \ldots, T. \end{cases}$$

Using this variable, we construct the following statistic, which we call the PS statistic

$$S_T^{ps} = \frac{1}{\tilde{\sigma}^2 T^2} y^{*\prime} MLL' M y^* = \frac{1}{\tilde{\sigma}^2 T^2} \sum_{j=1}^{T-1} \left(\sum_{t=1}^{j} \tilde{x}_t^* \right)^2,$$

where $\tilde{\sigma}^2$ is defined by (4) and $\tilde{x}_t^*$ are the regression residuals of y_t^* on z_t,

$$\tilde{x}_t^* = y_t^* - z_t' \left(\sum_{t=1}^{T} z_t z_t' \right)^{-1} \sum_{t=1}^{T} z_t y_t^*.$$

The key point is that the test statistic is decomposed into two terms, and the appropriate order is imposed on each term. Note that the test statistics for Cases 1 and 2 cannot be decomposed into two terms as in Cases 0 and 3. (This is why we cannot construct the test statistic whose limiting distribution is independent of the break point for Cases 1 and 2.)

The following theorem gives the limiting distribution of the PS statistic and its characteristic function.

Theorem 2. Consider model (1). *For Cases* 0 *and* 3, *under a sequence of local alternatives*, H_1,

$$S_T^{ps} \xrightarrow{\mathrm{d}} G(B_1; c^2\omega^2/\alpha^2) + G(B_2; c^2(1-\omega)^2/\alpha^2)$$

and its characteristic function is expressed as

$$\begin{aligned} \phi(\theta; c) = {} & [D(\mathrm{i}\theta + \sqrt{-\theta^2 + 2\mathrm{i}c^2\omega^2\theta})D(\mathrm{i}\theta - \sqrt{-\theta^2 + 2\mathrm{i}c^2\omega^2\theta})]^{-1/2} \\ & [D(\mathrm{i}\theta + \sqrt{-\theta^2 + 2\mathrm{i}c^2(1-\omega)^2\theta}) \\ & D(\mathrm{i}\theta - \sqrt{-\theta^2 + 2\mathrm{i}c^2(1-\omega)^2\theta})]^{-1/2}, \end{aligned} \tag{14}$$

where $G(B_1)$, $G(B_2)$ *and* $D(\lambda)$ *are defined as in Theorem 1* (ia) *and* (ib) *for Cases* 0 *and* 3, *respectively.*

Remark 5. Although the above limiting distribution depends on the value of ω under H_1, we have, for $c=0$,

$$\begin{aligned} S_T^{ps} \xrightarrow{\mathrm{d}} & \left(\int_0^1 B_1(r)^2\,\mathrm{d}r - X(B_1)'\Lambda^{-1}X(B_1) \right) \\ & + \left(\int_0^1 B_2(r)^2\,\mathrm{d}r - X(B_2)'\Lambda^{-1}X(B_2) \right) \end{aligned}$$

and $\phi(\theta, 0) = [D(2\mathrm{i}\theta)]^{-1}$, so that the null distribution does not depend on the break point.

Remark 6. Note that when $\omega = 0.5$, $T/T_B = T/(T - T_B) = 1/2$ so that $S_T^{PS} = (1/4)S_T$. Thus, the PS test is equivalent to the LM test in finite samples as well as in the limit when the break point is located at the center of the sample.

Table 2
Percent points of the null distribution of the PS test

	0.01	0.05	0.1	0.5	0.9	0.95	0.99
Case 0	0.07883	0.10942	0.13222	0.27757	0.60704	0.74752	1.07366
Case 3	0.04912	0.06265	0.07184	0.12087	0.21067	0.24654	0.32862

As in the case of the LM test, we can calculate the percentiles of the PS test under H_0 by numerical integration using the inversion formula (13). Table 2 reports each percentage point of the PS test for Cases 0 and 3.

Although the null distribution does not depend on the break point, it depends on ω under a sequence of local alternatives as shown by Theorem 2, so that the power depends on the location of the break point. Figs. 3a and b show the limiting power functions of the PS tests for Cases 0 and 3 with $\alpha = 1$. Again, as the characteristic function is symmetric around $\omega = 0.5$, we consider only the cases where $\omega \leqslant 0.5$. The relations among power functions are very similar to those of the LM test. That is, the power function corresponding to the smaller value of ω dominates that corresponding to the larger value of ω. However, the differences among the values of ω are not as large as for the LM test for both Cases 0 and 3.

Now we have two test statistics, S_T and S_T^{ps}, for Cases 0 and 3. Our interest now is in the difference of the powers of their limiting distributions, and whether one dominates the other with respect to power. Fig. 4 depicts the limiting power functions of the LM and PS tests for Cases 0 and 3 when $\omega = 0.3$. From the figure, we can see that the power of the LM test dominates that of the PS test for small values of c but that this relation is reversed when c increases, although the difference between their powers is slight. As the LM test is LBI, the dominance of the LM test local to the null can be seen as a theoretical result. Note that, as discussed in Remark 6, their power functions are exactly the same when $\omega = 0.5$, and so the PS test can be seen as the LBI test in such a case.

3.3. The innovational outlier model

Until now, we have investigated the additive outlier model in which the structural change affects the observation only at one time. Here, we discuss the innovational outlier model. That is, we consider the case when the shock is experienced gradually.

Let us consider the following model:

$$y_t = z_{1t}'\beta_1 + \psi(B)(z_{2t}'\beta_2) + x_t, \tag{15}$$

where $z_{1t} = 1$ or $[1, t/T]'$, $z_{2t} = DU_t$, DT_t, or $[DU_t, DT_t]'$ according to Cases 0–3, $\psi(B) = 1 + \psi_1 B + \cdots + \psi_m B^m$ is an mth order lag polynomial, and x_t is

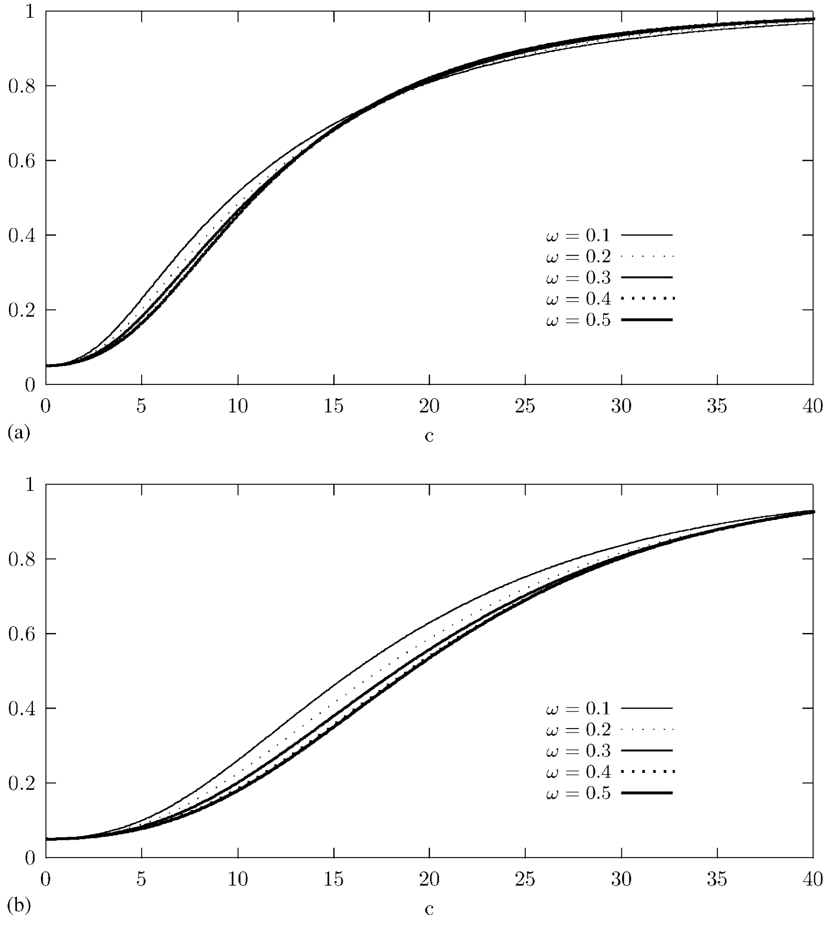


Fig. 3. (a) The limiting powers (Case 0: PS), (b) (Case 3: PS).

defined as in model (1). By introducing the lag polynomial $\psi(B)$, the shock of the structural change affects y_t gradually with lags.

Note that model (15) has the same expression as (1) with $z_t = [z_{1t}', z_{2t}', z_{2t-1}', \ldots, z_{2t-m}']'$ and $\beta = [\beta_1', \beta_2', \psi_1\beta_2', \ldots, \psi_m\beta_2']'$. Then, the LM test statistic can be constructed in the same way as (3).

To consider the limiting distribution of the test statistic we investigate $\tilde{x}_t$, the regression residuals of y_t on z_t. Note that we can write

$$\psi(B)DU_t = \eta_0 DU_t + \mathrm{d}(t, T_B)\check{\eta},$$

$$\psi(B)DT_t = \gamma_0 DU_t + \gamma_1 DT_t + \mathrm{d}(t, T_B)\check{\gamma},$$

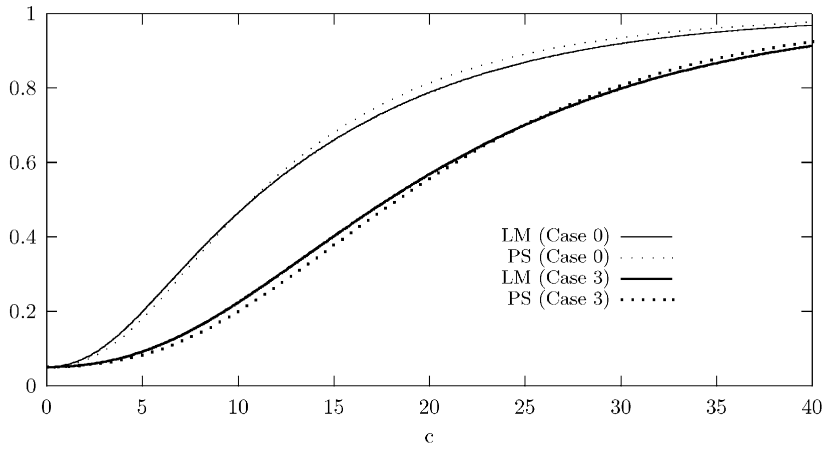


Fig. 4. The limiting powers (Cases 0 and 3: $\omega = 0.3$).

where η_0, $\check{\eta} = [\eta_1, \ldots, \eta_m]'$, γ_0, γ_1, and $\check{\gamma} = [\gamma_2, \ldots, \gamma_{m+1}]'$ are implicitly defined, and $\mathrm{d}(t, T_B) = [D(T_B)_t, \ldots, D(T_B)_{t-m}]$ with $D(T_B)_t = 1(t = T_B + 1)$. Some elements of $\check{\eta}$ and $\check{\gamma}$ might be zero. Then, $\tilde{x}_t$ is equivalent to the regression residuals of y_t on $\check{z}_t$, where

$$\check{z}_t = \begin{cases} [1, DU_t, \mathrm{d}(t, T_B)]' & \text{for Case 0,} \\ [1, t, DU_t, \mathrm{d}(t, T_B)]' & \text{for Case 1,} \\ [1, t, DU_t, DT_t, \mathrm{d}(t, T_B)]' & \text{for Case 2,} \\ [1, t, DU_t, DT_t, \mathrm{d}(t, T_B)]' & \text{for Case 3.} \end{cases}$$

However, because $\mathrm{d}(t, T_B)$ is asymptotically negligible, $\tilde{x}_t$ can be seen as regression residuals of y_t on $\check{z}_t = [1, DU_t]$, $[1, t/T, DU_t]$, $[1, t/T, DU_t, DT_t]$, and $[1, t/T, DU_t, DT_t]$ for Cases 0, 1, 2 and 3, respectively. Then, for Cases 0, 1 and 3, the limiting distributions of the test statistics in the innovational outlier model are the same as those in the additive outlier model, whereas, for Case 2, the limiting distribution is the same as in Case 3. Then, if we investigate the time series with the innovational outlier model, we can refer to Tables 1a, b, c, and d for Cases 0, 1, 2, and 3, respectively.

4. Testing for stationarity with an unknown break point

In this section, we consider the case when the break point is unknown. In such a case, we have to estimate the break point and construct the test statistic using the estimated break point.

The estimation of the break point has been considered in the literature, and we use the least-squares method found in Bai (1994, 1997) and Nunes et al. (1995). Let $SSR(k)$ denote the sum of squared residuals from the regression of y_t on z_t assuming that a structural change occurred at time k. Then the least-squares estimator of the break point denoted by T_B^* is defined by

$$T_B^* = \arg \min_{1 \leqslant k \leqslant T} SSR(k).$$

According to Bai (1994, 1997) and Nunes et al. (1995), the estimated fraction of the pre-break point observations, defined by $\omega^* = T_B^*/T$, is consistent when x_t is an $I(0)$ stochastic process. Moreover, Bai (1994, 1997) proved that $T(\omega^* - \omega) = T_B^* - T_B$ is $O_p(1)$. Using this result, we have the following corollary.

Corollary 1. Assume that $\{v_t, \varepsilon_t\}'$ *is normally distributed and construct the test statistics* S_T *and* S_T^{ps} *using the estimated break point* T_B^* *instead of* T_B. *Then, under the local alternative* H_1, *both test statistics have the same limiting distribution as in Theorems 1 and 2.*

The assumption of normality is sufficient for $x_t = u_t + \gamma_t$ to satisfy the assumption A6b in Bai (1997). Corollary 1 indicates that the limiting properties investigated in Section 3 hold even if we use the estimated break point.

However, the local asymptotic theory may be seen as rather theoretical, and we should also consider the fixed alternative. As discussed in Remark 1, the test statistic is consistent against the fixed alternative when the break point is known. This is not obvious when it is estimated because ω^* may not be consistent under the fixed alternative. For example, we can show that ω^* does not converge in probability to ω for Case 0 when $\rho > 0$ is fixed, and in such a case we will use the wrong break date. We are also interested in the behavior of the test statistic when we pre-specify the break date, by using, for example, visual inspection of the data, as well as when there really is no break.

Then let us consider the case when we use the pre-specified or estimated break date T_B^*, where T_B^* is fixed or of order at most T but ω^* does not necessarily converge in probability to ω, and define z_t^* in the same way as z_t with the break date T_B^*. Note that $\{\gamma_t\}$ dominates the behavior of $y_t = z_t'\beta + \gamma_t + u_t$, as $\{\gamma_t\}$ is an $I(1)$ process under the fixed alternative, while the other two terms are $O_p(1)$. Then we have

$$\sum_{t=1}^{j} \tilde{x} = \sum_{t=1}^{j} y_t - \sum_{t=1}^{j} z_t^{*\prime} \left(\sum_{t=1}^{T} z_t^* z_t^{*\prime} \right)^{-1} \sum_{t=1}^{T} z_t^* y_t = O_p(T^{3/2}),$$

so that $\sum_{j=1}^{T}(\sum_{t=1}^{j}\tilde{x}_t)^2$ is $O_p(T^4)$ as the known break point case. Similarly, we can deduce that $\tilde{\sigma}^2$ is $O_p(\ell T)$. Then, the test statistic has the same order, $O_p(T/\ell)$, as the case when $T_B^* = T_B$. This shows that the LM test statistic is consistent even when the break point is misspecified under the fixed alternative, although the local asymptotic theory discussed in the previous section does not hold in this case. Similarly, consistency of the PS test can also be shown.

We also note that our analysis is established under the assumption of the existence of a structural change. Then, in practice, the test for a structural change may be useful before our test is conducted. See, for example, Andrews (1993), Andrews and Ploberger (1994), and Vogelsang (1997) for tests of a structural change.

5. Finite sample properties

In this section, we investigate the finite sample behavior of the LM test statistic S_T and the PS test statistic S_T^{ps} for the sample sizes $T = 100$ and 200. We consider the following data generating process (DGP):

$$y_t = z_t'\beta + \gamma_t + u_t, \quad \gamma_t = \gamma_{t-1} + \varepsilon_t, \quad u_t = au_{t-1} + v_t, \tag{16}$$

where $\varepsilon_t \sim NID(0,\rho)$, $v_t \sim NID(0,1)$, $\{\varepsilon_t\}$ and $\{v_t\}$ are independent, $\gamma_0 = 0$, $u_0 = 0$, and $z_t'\beta = 1 + DU_t$, $1 + DU_t + 0.2t$, $1 + 0.2t + 0.02(t - T_B)$, and $1 + DU_t + 0.2t + 0.02(t - T_B)$ for Cases 0, 1, 2 and 3, respectively. The size of the test depends on α and ω whereas the power is affected by ρ as well as those parameters. We set $a = 0$, ± 0.2, ± 0.5 and ± 0.8, $\omega = 0.2$, 0.5 and 0.8, and $\rho = 0.01$ and 1. Simulations are conducted with the known break point, T_B, as well as with the estimated break point, T_B^*, which is estimated by the least-squares method as explained in Section 4. The number of replications is 1000 in all experiments, performed by the GAUSS matrix programming language.

For the construction of $\tilde{\sigma}^2$, we use the Bartlett window defined by $w(\mathrm{i},\ell) = 1 - \mathrm{i}/(\ell + 1)$. We also have to choose the lag truncation number ℓ, which affects the finite sample properties of the test both under the null and under the alternative. We consider two types of lag truncation number. One depends only on the sample size T: $\ell 4 = [4(T/100)^{1/4}]$ and $\ell 12 = [12(T/100)^{1/4}]$, as in KPSS (1992b). The other is the data dependent selection rule, which is a slight modification of the automatic bandwidth parameter proposed in Andrews (1991). Here we should note that the Andrews' automatic bandwidth parameter, ℓ_A, cannot be applied to our test statistic. For example, ℓ_A

for an AR(1) model is defined by

$$\ell_{\mathrm{A}} = 1.1447\left\{\frac{4\hat{a}^2 T}{(1+\hat{a})^2(1-\hat{a})^2}\right\}^{1/3}, \tag{17}$$

where $\hat{a}$ is an estimate of the autoregressive parameter. Under the null hypothesis, $\hat{a}$ converges in probability to a and $4\hat{a}^2/(1+\hat{a})^2/(1-\hat{a})^2$ to $4a^2/(1+a)^2/(1-a)^2 > 0$ as $|a| < 1$. Then, ℓ_{A} is of order $T^{1/3}$, that satisfies the condition $\ell = \mathrm{o_p}(T^{1/2})$. On the other hand, the process has a unit root under the alternative and then $(1-\hat{a})$ is of order T^{-1}. As a result, the order of ℓ_{A} becomes T. Since the test statistic S_T is $\mathrm{O_p}(T/\ell)$ under the fixed alternative as discussed in Remark 1 and Section 4, S_T is not consistent in the usual sense if we use the automatic bandwidth parameter. Note that for the same reason, we cannot use the automatic bandwidth parameter for the KPSS test.

One possibility for making the test statistic consistent is to constrain ℓ_{A} to diverge at a slower rate than T by introducing the upper bound of order less than T. As a bound, we consider the optimal fixed bandwidth parameter in Andrews (1991), which is defined in the same way as (17) with $\hat{a}$ pre-specified, not estimated from the data. Then we use the following data dependent bandwidth parameter:

$$\ell_{\mathrm{A}}k = \min\left(1.1447\left\{\frac{4\hat{a}^2 T}{(1+\hat{a})^2(1-\hat{a})^2}\right\}^{1/3}, 1.1447\left\{\frac{k^2 T}{(1+k)^2(1-k)^2}\right\}^{1/3}\right).$$

We consider $k = 0.6$, 0.7, 0.8, and 0.9, as many economic time series have positive autocorrelations. Note that because $\ell_{\mathrm{A}}k$ is $\mathrm{O_p}(T^{1/3})$ both under the null and under the alternative, $\tilde{\sigma}^2$ with $\ell_{\mathrm{A}}k$ is consistent under the null and S_T diverges under the alternative.

Tables 3a and b reports the empirical sizes of the LM and PS tests. Entries in the upper row for each ω are the empirical sizes when the break point is known, whereas those in the unknown break point case are tabulated in the lower row with parentheses. As we used the upper 5% point as the critical value, the nominal size of the test is 0.05. The results of the simulation are summarized as follows:

(i) The size of the LM test is much affected by the persistence of the stationary component, a, and the lag truncation number, ℓ. As a whole, there is tendency for over-rejection when the value of a goes to 1 and for under-rejection when a is a negative value.

(ii) The larger the sample size becomes, the closer is the empirical size to the nominal one.

(iii) The test with $\ell 4$ has a large size distortion when a is 0.8. The test with $\ell 12$ has a result similar to that with $\ell_{\mathrm{A}}0.7$. As a whole, tests with $\ell_{\mathrm{A}}0.8$ and $\ell_{\mathrm{A}}0.9$ have closer empirical sizes to the nominal one than the others.

Table 3
The size of the tests

ω	a	$T=100$						$T=200$					
		$\ell 4$	$\ell 12$	$\ell_A 0.6$	$\ell_A 0.7$	$\ell_A 0.8$	$\ell_A 0.9$	$\ell 4$	$\ell 12$	$\ell_A 0.6$	$\ell_A 0.7$	$\ell_A 0.8$	$\ell_A 0.9$
(a) (Case 0: LM)													
0.2	0.8	0.253	0.107	0.147	0.118	0.103	0.099	0.287	0.110	0.138	0.113	0.098	0.098
		(0.211)	(0.085)	(0.112)	(0.098)	(0.087)	(0.087)	(0.230)	(0.082)	(0.104)	(0.084)	(0.077)	(0.077)
	0.5	0.098	0.068	0.083	0.083	0.083	0.083	0.091	0.072	0.069	0.069	0.069	0.069
		(0.076)	(0.059)	(0.064)	(0.064)	(0.064)	(0.064)	(0.075)	(0.058)	(0.062)	(0.062)	(0.062)	(0.062)
	0.2	0.057	0.055	0.059	0.059	0.059	0.059	0.049	0.055	0.053	0.053	0.053	0.053
		(0.046)	(0.046)	(0.047)	(0.047)	(0.047)	(0.047)	(0.048)	(0.048)	(0.047)	(0.047)	(0.047)	(0.047)
	0	0.038	0.046	0.047	0.047	0.047	0.047	0.040	0.050	0.041	0.041	0.041	0.041
		(0.034)	(0.035)	(0.038)	(0.038)	(0.038)	(0.038)	(0.042)	(0.045)	(0.042)	(0.042)	(0.042)	(0.042)
	−0.2	0.032	0.036	0.030	0.030	0.030	0.030	0.031	0.040	0.032	0.032	0.032	0.032
		(0.030)	(0.029)	(0.029)	(0.029)	(0.029)	(0.029)	(0.031)	(0.040)	(0.031)	(0.031)	(0.031)	(0.031)
	−0.5	0.021	0.027	0.027	0.027	0.027	0.027	0.017	0.030	0.026	0.026	0.026	0.026
		(0.019)	(0.027)	(0.023)	(0.023)	(0.023)	(0.023)	(0.017)	(0.031)	(0.022)	(0.022)	(0.022)	(0.022)
	−0.8	0.001	0.022	0.014	0.019	0.023	0.025	0.006	0.016	0.011	0.021	0.020	0.022
		(0.014)	(0.027)	(0.026)	(0.027)	(0.030)	(0.032)	(0.005)	(0.020)	(0.013)	(0.020)	(0.022)	(0.024)
0.5	0.8	0.323	0.064	0.139	0.094	0.066	0.063	0.365	0.075	0.121	0.082	0.053	0.053
		(0.193)	(0.074)	(0.108)	(0.088)	(0.080)	(0.081)	(0.244)	(0.083)	(0.109)	(0.086)	(0.072)	(0.072)
	0.5	0.113	0.044	0.082	0.082	0.082	0.082	0.096	0.038	0.057	0.057	0.057	0.057
		(0.068)	(0.037)	(0.054)	(0.054)	(0.054)	(0.054)	(0.065)	(0.036)	(0.045)	(0.045)	(0.045)	(0.045)
	0.2	0.054	0.041	0.068	0.068	0.068	0.068	0.045	0.031	0.050	0.050	0.050	0.050
		(0.049)	(0.036)	(0.059)	(0.059)	(0.059)	(0.059)	(0.029)	(0.026)	(0.035)	(0.035)	(0.035)	(0.035)
	0	0.038	0.035	0.046	0.046	0.046	0.046	0.034	0.031	0.031	0.031	0.031	0.031
		(0.036)	(0.029)	(0.036)	(0.036)	(0.036)	(0.036)	(0.024)	(0.020)	(0.024)	(0.024)	(0.024)	(0.024)
	−0.2	0.029	0.032	0.030	0.030	0.030	0.030	0.022	0.030	0.021	0.021	0.021	0.021
		(0.028)	(0.026)	(0.029)	(0.029)	(0.029)	(0.029)	(0.015)	(0.017)	(0.015)	(0.015)	(0.015)	(0.015)
	−0.5	0.018	0.027	0.020	0.020	0.020	0.020	0.015	0.025	0.018	0.018	0.018	0.018
		(0.022)	(0.024)	(0.022)	(0.022)	(0.022)	(0.022)	(0.009)	(0.014)	(0.013)	(0.013)	(0.013)	(0.013)

	−0.8	0.003	0.023	0.013	0.023	0.029	0.034	0.001	0.018	0.009	0.016	0.020	0.022
		(0.010)	(0.020)	(0.015)	(0.020)	(0.024)	(0.030)	(0.000)	(0.007)	(0.003)	(0.006)	(0.009)	(0.012)
0.8	0.8	0.291	0.122	0.168	0.137	0.118	0.114	0.266	0.101	0.128	0.107	0.089	0.089
		(0.190)	(0.071)	(0.094)	(0.084)	(0.081)	(0.082)	(0.222)	(0.076)	(0.096)	(0.079)	(0.072)	(0.071)
	0.5	0.101	0.056	0.087	0.087	0.087	0.087	0.095	0.054	0.069	0.069	0.069	0.069
		(0.074)	(0.054)	(0.063)	(0.063)	(0.063)	(0.063)	(0.076)	(0.050)	(0.064)	(0.064)	(0.064)	(0.064)
	0.2	0.058	0.039	0.057	0.057	0.057	0.057	0.049	0.043	0.051	0.051	0.051	0.051
		(0.050)	(0.048)	(0.052)	(0.052)	(0.052)	(0.052)	(0.050)	(0.046)	(0.052)	(0.052)	(0.052)	(0.052)
	0	0.039	0.034	0.037	0.037	0.037	0.037	0.039	0.038	0.039	0.039	0.039	0.039
		(0.036)	(0.039)	(0.034)	(0.034)	(0.034)	(0.034)	(0.037)	(0.043)	(0.040)	(0.040)	(0.040)	(0.040)
	−0.2	0.028	0.031	0.026	0.026	0.026	0.026	0.029	0.037	0.027	0.027	0.027	0.027
		(0.030)	(0.033)	(0.030)	(0.030)	(0.030)	(0.030)	(0.029)	(0.034)	(0.027)	(0.027)	(0.027)	(0.027)
	−0.5	0.015	0.023	0.018	0.018	0.018	0.018	0.019	0.029	0.020	0.020	0.020	0.020
		(0.018)	(0.026)	(0.022)	(0.022)	(0.022)	(0.022)	(0.019)	(0.027)	(0.023)	(0.023)	(0.023)	(0.023)
	−0.8	0.001	0.014	0.005	0.012	0.015	0.018	0.001	0.010	0.010	0.011	0.012	0.013
		(0.013)	(0.027)	(0.018)	(0.023)	(0.028)	(0.029)	(0.004)	(0.014)	(0.012)	(0.018)	(0.024)	(0.024)
(b) (Case 0: PS)													
0.2	0.8	0.195	0.037	0.079	0.053	0.045	0.045	0.301	0.081	0.106	0.086	0.066	0.065
		(0.178)	(0.061)	(0.074)	(0.061)	(0.059)	(0.059)	(0.227)	(0.064)	(0.086)	(0.071)	(0.064)	(0.066)
	0.5	0.081	0.064	0.068	0.068	0.068	0.068	0.104	0.059	0.067	0.067	0.067	0.067
		(0.056)	(0.057)	(0.051)	(0.051)	(0.051)	(0.051)	(0.080)	(0.051)	(0.063)	(0.063)	(0.063)	(0.063)
	0.2	0.051	0.069	0.065	0.065	0.065	0.065	0.054	0.052	0.057	0.057	0.057	0.057
		(0.034)	(0.063)	(0.040)	(0.040)	(0.040)	(0.040)	(0.051)	(0.048)	(0.050)	(0.050)	(0.050)	(0.050)
	0	0.044	0.066	0.054	0.054	0.054	0.054	0.041	0.048	0.038	0.038	0.038	0.038
		(0.030)	(0.060)	(0.039)	(0.039)	(0.039)	(0.039)	(0.035)	(0.048)	(0.040)	(0.040)	(0.040)	(0.040)
	−0.2	0.034	0.068	0.035	0.035	0.035	0.035	0.028	0.045	0.027	0.027	0.027	0.027
		(0.032)	(0.064)	(0.033)	(0.033)	(0.033)	(0.033)	(0.023)	(0.043)	(0.023)	(0.023)	(0.023)	(0.023)
	−0.5	0.020	0.069	0.032	0.032	0.032	0.032	0.016	0.040	0.022	0.022	0.022	0.022
		(0.024)	(0.065)	(0.036)	(0.035)	(0.035)	(0.035)	(0.015)	(0.035)	(0.025)	(0.025)	(0.025)	(0.025)
	−0.8	0.005	0.076	0.029	0.053	0.102	0.126	0.001	0.028	0.018	0.034	0.040	0.044
		(0.011)	(0.072)	(0.031)	(0.054)	(0.091)	(0.122)	(0.004)	(0.024)	(0.016)	(0.027)	(0.030)	(0.032)
0.5	0.8	0.323	0.064	0.139	0.094	0.066	0.063	0.365	0.075	0.121	0.082	0.053	0.053

Table 3. (*continued*)

ω	a	$T=100$						$T=200$					
		$\ell 4$	$\ell 12$	$\ell_{\mathrm{A}}0.6$	$\ell_{\mathrm{A}}0.7$	$\ell_{\mathrm{A}}0.8$	$\ell_{\mathrm{A}}0.9$	$\ell 4$	$\ell 12$	$\ell_{\mathrm{A}}0.6$	$\ell_{\mathrm{A}}0.7$	$\ell_{\mathrm{A}}0.8$	$\ell_{\mathrm{A}}0.9$
		(0.186)	(0.060)	(0.084)	(0.070)	(0.065)	(0.066)	(0.270)	(0.077)	(0.104)	(0.084)	(0.072)	(0.071)
	0.5	0.113	0.044	0.082	0.082	0.082	0.082	0.096	0.038	0.057	0.057	0.057	0.057
		(0.069)	(0.041)	(0.053)	(0.053)	(0.053)	(0.053)	(0.079)	(0.044)	(0.051)	(0.051)	(0.051)	(0.051)
	0.2	0.054	0.041	0.068	0.068	0.068	0.068	0.045	0.031	0.050	0.050	0.050	0.050
		(0.048)	(0.038)	(0.056)	(0.056)	(0.056)	(0.056)	(0.034)	(0.028)	(0.035)	(0.035)	(0.035)	(0.035)
	0	0.038	0.035	0.046	0.046	0.046	0.046	0.034	0.031	0.031	0.031	0.031	0.031
		(0.030)	(0.036)	(0.040)	(0.040)	(0.040)	(0.040)	(0.023)	(0.020)	(0.024)	(0.024)	(0.024)	(0.024)
	−0.2	0.029	0.032	0.030	0.030	0.030	0.030	0.022	0.030	0.021	0.021	0.021	0.021
		(0.026)	(0.029)	(0.027)	(0.027)	(0.027)	(0.027)	(0.016)	(0.018)	(0.016)	(0.016)	(0.016)	(0.016)
	−0.5	0.018	0.027	0.020	0.020	0.020	0.020	0.015	0.025	0.018	0.018	0.018	0.018
		(0.020)	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.008)	(0.014)	(0.012)	(0.012)	(0.012)	(0.012)
	−0.8	0.003	0.023	0.013	0.023	0.029	0.034	0.001	0.018	0.009	0.016	0.020	0.022
		(0.009)	(0.018)	(0.013)	(0.020)	(0.029)	(0.036)	(0.000)	(0.007)	(0.004)	(0.007)	(0.010)	(0.014)
0.8	0.8	0.257	0.061	0.098	0.064	0.059	0.059	0.296	0.074	0.108	0.080	0.059	0.057
		(0.164)	(0.061)	(0.069)	(0.053)	(0.050)	(0.050)	(0.227)	(0.061)	(0.088)	(0.066)	(0.055)	(0.056)
	0.5	0.088	0.064	0.067	0.066	0.066	0.066	0.099	0.047	0.065	0.065	0.065	0.065
		(0.060)	(0.056)	(0.049)	(0.049)	(0.049)	(0.049)	(0.081)	(0.037)	(0.058)	(0.058)	(0.058)	(0.058)
	0.2	0.049	0.069	0.052	0.052	0.052	0.052	0.051	0.037	0.056	0.056	0.056	0.056
		(0.047)	(0.064)	(0.050)	(0.050)	(0.050)	(0.050)	(0.041)	(0.030)	(0.047)	(0.047)	(0.047)	(0.047)
	0	0.037	0.073	0.041	0.041	0.041	0.041	0.038	0.037	0.046	0.046	0.046	0.046
		(0.046)	(0.072)	(0.039)	(0.039)	(0.039)	(0.039)	(0.033)	(0.030)	(0.039)	(0.039)	(0.039)	(0.039)
	−0.2	0.034	0.070	0.036	0.036	0.036	0.036	0.031	0.037	0.027	0.027	0.027	0.027
		(0.037)	(0.074)	(0.033)	(0.033)	(0.033)	(0.033)	(0.026)	(0.027)	(0.024)	(0.024)	(0.024)	(0.024)
	−0.5	0.026	0.071	0.042	0.043	0.043	0.043	0.017	0.028	0.024	0.024	0.024	0.024
		(0.025)	(0.079)	(0.039)	(0.038)	(0.038)	(0.038)	(0.014)	(0.023)	(0.017)	(0.017)	(0.017)	(0.017)
	−0.8	0.008	0.100	0.046	0.077	0.122	0.150	0.001	0.024	0.018	0.022	0.030	0.033
		(0.013)	(0.096)	(0.045)	(0.076)	(0.108)	(0.144)	(0.004)	(0.024)	(0.014)	(0.025)	(0.038)	(0.040)

(iv) The PS test seems to have better finite sample properties under the null hypothesis than the LM test.
(v) Tests with an estimated break point have an empirical size closer to the nominal one than those with a known break point.

The results for Cases 1–3 are similar to Case 0, but the rejection frequencies for the LM test with a known break point for these cases tend to be larger than for Case 0 when $a = 0.8$. For example, the empirical sizes with $\ell_A 0.8$ when $a = 0.8$ and $\omega = 0.2$ are 0.103, 0.130, 0.162 and 0.200 for Cases 0, 1, 2 and 3, respectively, when $T = 100$ and 0.098, 0.102, 0.114 and 0.147 for $T = 200$. We also note that the PS statistics with $\ell_A 0.7$, $\ell_A 0.8$ and $\ell_A 0.9$ for Case 3 have a large size distortion when $a = -0.8$ and $\omega = 0.2$ or $\omega = 0.8$. To save space, we do not tabulate these results but they are available upon request.

As mentioned above, the test tends to reject too frequently when a is 0.8, and so we investigated the case when a is closer to 1, such as 0.9. The empirical sizes with $\ell_A 0.8$ when the break point is known and $\omega = 0.2$ are 0.159, 0.030, 0.197, 0.217, 0.282, 0.037 for Cases 0, 0PS, 1, 2, 3 and 3PS for $T = 100$ and 0.171, 0.099, 0.176, 0.182, 0.226 and 0.073 when $T = 200$. There still remains a tendency for over-rejection when $T = 200$. This means that even when the null of trend stationarity is rejected with a known break point, we cannot conclude with certainty that the process is $I(1)$.

On the other hand, the empirical sizes when the break point is estimated are 0.097, 0.045, 0.088, 0.063, 0.156 and 0.120 for $T = 100$ and 0.091, 0.075, 0.070, 0.071, 0.108 and 0.070 for $T = 200$ for the corresponding cases. Then the tests with the estimated break point have a less severe size distortion than those with a known break point.

Table 4a and b reports the empirical power of the tests. The results of the simulation are summarized as follows:

(i) In most cases, power increases when the sample size becomes large.
(ii) The tests with $\ell_A 0.8$ and $\ell_A 0.9$ tend to be less powerful than the others.
(iii) As a becomes large, the power tends to be lower as indicated in the previous section.
(iv) It seems that the LM test is more powerful than the PS test.
(v) When the break point is estimated, power decreases considerably in some cases compared with the known break point case.

The results of the other cases are similar to Case 0 and to save space we do not report them.

Although it is difficult to conclude from our simulations which selection rules for the lag truncation number are best, considering both size and power properties, the test with $\ell_A 0.7$ and $\ell_A 0.8$ may have better finite sample properties than the others.

Table 4
The power of the tests (Case 0: LM).

		$T=100$						$T=200$					
ρ	a	$\ell 4$	$\ell 12$	$\ell_A 0.6$	$\ell_A 0.7$	$\ell_A 0.8$	$\ell_A 0.9$	$\ell 4$	$\ell 12$	$\ell_A 0.6$	$\ell_A 0.7$	$\ell_A 0.8$	$\ell_A 0.9$
(a) (Case 0: LM)													
0.01	0.8	0.347	0.171	0.221	0.190	0.162	0.157	0.470	0.263	0.302	0.271	0.239	0.231
		(0.199)	(0.087)	(0.111)	(0.097)	(0.093)	(0.094)	(0.286)	(0.113)	(0.138)	(0.119)	(0.108)	(0.108)
	0.5	0.294	0.218	0.261	0.258	0.258	0.258	0.566	0.458	0.507	0.506	0.506	0.506
		(0.138)	(0.097)	(0.124)	(0.124)	(0.124)	(0.124)	(0.308)	(0.180)	(0.235)	(0.234)	(0.234)	(0.234)
	0.2	0.369	0.295	0.383	0.383	0.383	0.383	0.665	0.572	0.655	0.655	0.655	0.655
		(0.197)	(0.137)	(0.203)	(0.203)	(0.203)	(0.203)	(0.444)	(0.307)	(0.439)	(0.439)	(0.439)	(0.439)
	0	0.431	0.342	0.453	0.453	0.453	0.453	0.709	0.605	0.739	0.739	0.739	0.739
		(0.239)	(0.166)	(0.264)	(0.264)	(0.264)	(0.264)	(0.511)	(0.359)	(0.558)	(0.558)	(0.558)	(0.558)
	−0.2	0.475	0.378	0.499	0.499	0.499	0.499	0.749	0.632	0.782	0.782	0.782	0.782
		(0.287)	(0.174)	(0.311)	(0.311)	(0.311)	(0.311)	(0.565)	(0.392)	(0.609)	(0.609)	(0.609)	(0.609)
	−0.5	0.505	0.412	0.500	0.500	0.500	0.500	0.778	0.662	0.763	0.763	0.763	0.763
		(0.320)	(0.208)	(0.322)	(0.322)	(0.322)	(0.322)	(0.616)	(0.415)	(0.578)	(0.578)	(0.578)	(0.578)
	−0.8	0.454	0.426	0.458	0.446	0.439	0.441	0.743	0.657	0.701	0.692	0.681	0.680
		(0.261)	(0.217)	(0.260)	(0.254)	(0.250)	(0.252)	(0.552)	(0.419)	(0.473)	(0.450)	(0.440)	(0.440)
1	0.8	0.753	0.540	0.627	0.568	0.498	0.368	0.915	0.703	0.759	0.714	0.662	0.560
		(0.498)	(0.236)	(0.314)	(0.257)	(0.209)	(0.198)	(0.787)	(0.446)	(0.554)	(0.468)	(0.374)	(0.277)
	0.5	0.780	0.572	0.649	0.615	0.532	0.411	0.917	0.709	0.767	0.720	0.664	0.573
		(0.522)	(0.268)	(0.359)	(0.304)	(0.259)	(0.244)	(0.821)	(0.457)	(0.580)	(0.483)	(0.405)	(0.323)
	0.2	0.789	0.588	0.657	0.621	0.551	0.454	0.923	0.711	0.765	0.721	0.664	0.577
		(0.546)	(0.283)	(0.381)	(0.326)	(0.284)	(0.275)	(0.832)	(0.472)	(0.585)	(0.498)	(0.410)	(0.342)
	0	0.797	0.592	0.661	0.627	0.567	0.481	0.923	0.710	0.767	0.723	0.665	0.584
		(0.544)	(0.287)	(0.385)	(0.340)	(0.307)	(0.300)	(0.835)	(0.478)	(0.590)	(0.507)	(0.420)	(0.361)
	−0.2	0.803	0.592	0.669	0.636	0.587	0.516	0.923	0.711	0.768	0.723	0.667	0.596
		(0.548)	(0.289)	(0.398)	(0.361)	(0.332)	(0.326)	(0.839)	(0.484)	(0.591)	(0.514)	(0.435)	(0.393)
	−0.5	0.800	0.589	0.686	0.665	0.642	0.613	0.925	0.710	0.772	0.731	0.683	0.639
		(0.560)	(0.283)	(0.458)	(0.436)	(0.421)	(0.420)	(0.843)	(0.482)	(0.604)	(0.535)	(0.484)	(0.464)

	−0.8	0.794	0.587	0.791	0.784	0.779	0.775	0.922	0.711	0.807	0.790	0.777	0.770
		(0.548)	(0.289)	(0.661)	(0.658)	(0.656)	(0.656)	(0.845)	(0.485)	(0.726)	(0.708)	(0.695)	(0.692)
(b) (Case 0: PS)													
0.01	0.8	0.275	0.055	0.106	0.069	0.048	0.045	0.452	0.157	0.220	0.171	0.136	0.122
		(0.175)	(0.049)	(0.068)	(0.054)	(0.049)	(0.049)	(0.290)	(0.080)	(0.123)	(0.089)	(0.072)	(0.071)
	0.5	0.253	0.096	0.178	0.175	0.174	0.174	0.524	0.344	0.417	0.412	0.412	0.412
		(0.108)	(0.057)	(0.088)	(0.088)	(0.088)	(0.088)	(0.288)	(0.137)	(0.193)	(0.192)	(0.192)	(0.192)
	0.2	0.305	0.168	0.312	0.312	0.312	0.312	0.651	0.444	0.634	0.634	0.634	0.634
		(0.132)	(0.072)	(0.150)	(0.150)	(0.150)	(0.150)	(0.427)	(0.243)	(0.416)	(0.416)	(0.416)	(0.416)
	0	0.360	0.190	0.399	0.399	0.399	0.399	0.693	0.505	0.732	0.732	0.732	0.732
		(0.176)	(0.078)	(0.214)	(0.214)	(0.214)	(0.214)	(0.497)	(0.307)	(0.537)	(0.537)	(0.537)	(0.537)
	−0.2	0.404	0.213	0.432	0.432	0.432	0.432	0.737	0.533	0.777	0.777	0.777	0.777
		(0.213)	(0.096)	(0.251)	(0.251)	(0.251)	(0.251)	(0.551)	(0.335)	(0.602)	(0.602)	(0.602)	(0.602)
	−0.5	0.436	0.239	0.438	0.438	0.438	0.438	0.770	0.562	0.738	0.738	0.738	0.738
		(0.257)	(0.109)	(0.261)	(0.261)	(0.261)	(0.261)	(0.586)	(0.351)	(0.543)	(0.543)	(0.543)	(0.543)
	−0.8	0.373	0.241	0.343	0.319	0.304	0.303	0.724	0.569	0.651	0.621	0.601	0.600
		(0.194)	(0.114)	(0.160)	(0.151)	(0.152)	(0.158)	(0.531)	(0.359)	(0.439)	(0.408)	(0.396)	(0.395)
1	0.8	0.710	0.297	0.492	0.372	0.147	0.054	0.916	0.619	0.719	0.644	0.523	0.148
		(0.483)	(0.194)	(0.272)	(0.221)	(0.167)	(0.158)	(0.791)	(0.419)	(0.514)	(0.443)	(0.320)	(0.202)
	0.5	0.731	0.334	0.530	0.408	0.172	0.061	0.928	0.641	0.727	0.665	0.542	0.163
		(0.511)	(0.217)	(0.318)	(0.250)	(0.208)	(0.200)	(0.827)	(0.444)	(0.542)	(0.462)	(0.347)	(0.242)
	0.2	0.740	0.338	0.547	0.425	0.192	0.085	0.934	0.652	0.729	0.667	0.549	0.176
		(0.540)	(0.225)	(0.342)	(0.279)	(0.238)	(0.230)	(0.831)	(0.450)	(0.543)	(0.469)	(0.359)	(0.268)
	0	0.748	0.339	0.555	0.438	0.213	0.111	0.937	0.650	0.728	0.669	0.551	0.192
		(0.541)	(0.229)	(0.352)	(0.295)	(0.256)	(0.250)	(0.829)	(0.445)	(0.544)	(0.469)	(0.372)	(0.295)
	−0.2	0.749	0.338	0.557	0.450	0.240	0.138	0.935	0.648	0.727	0.670	0.556	0.225
		(0.549)	(0.235)	(0.376)	(0.326)	(0.303)	(0.299)	(0.828)	(0.452)	(0.546)	(0.480)	(0.388)	(0.334)
	−0.5	0.752	0.338	0.588	0.503	0.337	0.253	0.936	0.646	0.729	0.672	0.577	0.310
		(0.564)	(0.235)	(0.431)	(0.403)	(0.391)	(0.390)	(0.830)	(0.450)	(0.566)	(0.507)	(0.450)	(0.424)
	−0.8	0.751	0.334	0.735	0.708	0.644	0.601	0.935	0.646	0.772	0.743	0.703	0.618
		(0.555)	(0.237)	(0.652)	(0.650)	(0.649)	(0.649)	(0.835)	(0.454)	(0.699)	(0.682)	(0.671)	(0.667)

6. Empirical results

In this section, we apply the testing procedure developed in the previous section to the data series of Nelson and Plosser (1982). The Nelson–Plosser data have been used in various studies. In particular, the existence of a unit root is one of the most interesting issues, and has been analyzed in Perron (1997) and Zivot and Andrews (1992), assuming trend stationarity with a break under the alternative. They considered a structural change corresponding to Case 1, except for common-stock prices and real wages, for which Case 3 is assumed. Their results are very similar except for real wages; with the model corresponding to our Case 1, the unit root hypothesis is rejected for the five macroeconomic time series, real GNP, nominal GNP, industrial production, employment and nominal wages, weakly rejected (at 10% level) for real per capita GNP, and, with the model corresponding to our Case 3, the null of a unit root is rejected for common-stock prices.

Before applying our tests proposed in the previous section, we should test for the presence of a structural change in the Nelson–Plosser data. This is investigated in Vogelsang (1997), and according to Tables 1 [1] and 3 in Vogelsang (1997), there seems to be a tendency for a structural change under the null hypothesis except for employment and money stock. We apply our tests to 12 macroeconomic time series, except for these two series.

The results are tabulated in Table 5. The break point and its fraction of the sample size in the third and fourth columns are estimated by the least-squares method. For comparison, the results of the unit root test in Zivot and Andrews (1992) are tabulated in the column labeled Z&A and those of $t^*_{\alpha,\theta}(1)$ and $t^*_{\alpha,\gamma}(2)$ using *t–sig* in Perron (1997) are in the column labeled Perron. The model of Case 1 is used for all the series except for common-stock prices and real wages, to which the model of Case 3 is applied. As was seen in the previous section, tests with $\ell 4$ and $\ell_{\mathrm{A}}0.6$ tend to reject the null hypothesis too often and so we do not calculate the statistics for them. We also calculate the PS test statistic for common-stock prices and real wages. From the table, we can see that the unit root and stationarity tests provide consistent results for eight out of 12 time series; the null of a unit root is supported for velocity and interest rates whereas there seems to be no unit root in nominal GNP, industrial production, unemployment, nominal wages and common-stock prices. There is a weak tendency for a unit root for consumer prices. However, we should be careful in automatically concluding that velocity, interest rates and consumer prices have a unit root as the LM test for the null of trend

[1] Since our interest is whether or not there is a structural change under the null hypothesis, we use Table 1, not Table 2.

Table 5
Test for stationarity[a]

Series	T	T_B	ω	Z&A	Perron	$\ell 12$	$\ell_A 0.7$	$\ell_A 0.8$	$\ell_A 0.9$
Real GNP	62	1929	0.3387	[b]	[b]	0.0920[c]	0.0898	0.0937[c]	0.0937[c]
Nominal GNP	62	1929	0.3387	[b]	[b]	0.0766	0.0774	0.0774	0.0774
Real per capita GNP	62	1940	0.5161	[c]	[c]	0.1588[d]	0.1572[d]	0.1595[d]	0.1595[d]
Industrial production	111	1929	0.6306	[b]	[b]	0.0701	0.0716	0.0712	0.0712
Unemployment	81	1929	0.4938	—	—	0.0668	0.0631	0.0631	0.0631
GNP deflator	82	1929	0.5000			0.0945	0.0947	0.0957	0.0968
Consumer prices	111	1878	0.1712			0.0992[c]	0.0995[c]	0.1009[c]	0.1285[c]
Nominal wages	71	1930	0.4366	[d]	[b]	0.0901	0.0878	0.0901	0.0901
Velocity	102	1950	0.8039			0.1190[d]	0.1218[d]	0.1175[d]	0.1178[d]
Interest rate	71	1934	0.4930			0.1416[d]	0.1468[d]	0.1387[d]	0.1431[d]
Common-stock prices	100	1939	0.6900	[b]	[b]	0.0524	0.0455	0.0455	0.0455
(the PS test)						0.1661	0.1445	0.1445	0.1445
Real wages	71	1940	0.5775		[b]	0.0610[c]	0.0428	0.0428	0.0428
(the PS test)						0.2527[d]	0.1774	0.1774	0.1774

[a]The columns labeled by Z&A and Perron are the results of the unit root test in Zivot and Andrews (1992) and Perron (1997).
[b]Denotes significance at 1% level.
[c]Denotes significance at 5% level.
[d]Denotes significance at 10% level.

stationarity is oversized and a unit root test has low power when the process is highly persistent. In fact, the first order autocorrelations of velocity and consumer prices after detrending are 0.80 and 0.88, although the interest rate is not so very persistent; the first order autocorrelation is 0.68. Then, the former two series still show the possibility of trend stationarity with a break.

For the other series, test results are inconclusive; for real wages, the result of the trend stationarity test changes, depending on whether or not we use $\ell 12$. For the GNP deflator, we cannot reject both the null of a unit root and the null of trend stationarity, which may be due to low power of both test statistics when the roots are close to one. On the other hand, both hypotheses are rejected for real GNP and real per capita GNP. Several explanations may be considered. One reason is that, although there are no unit roots in the processes, the LM test for stationarity rejected the null hypothesis because of its property of over-rejection when the process is highly persistent. We may also consider that both the unit root and stationary processes with a break are not adequate. Further investigation may be required for these series, possibly trying models other than the simple $I(0)$ or $I(1)$ model.

7. Conclusion

In this paper, we have developed a testing procedure for the null hypothesis of stationarity with a break against nonstationarity. We proposed the LM test and the PS test, which does not depend on the location of the break point, ω, under the null hypothesis. The local limiting power was also investigated and the tests were shown to be consistent against the alternative of a unit root. The simulation experiment revealed that the proposed test tends to reject the null of trend stationarity too often when the process is highly persistent and the break point is known, while the test with the estimated break point does not have a large size distortion, but its power decreases. Although several testing procedures are proposed to test for the null of a unit root against stationarity with a break, our tests suppose the null of trend stationarity. They do not compete but complement each other to investigate the persistence of the time series.

Acknowledgements

I owe special thanks to Katsuto Tanaka and Taku Yamamoto. I am also grateful to two anonymous referees, associated editor, Koichi Maekawa and Mitsuru Nakagawa for useful comments. All errors are my responsibility. This research was partially supported by JSPS Research Fellowships for Young Scientists and by the Ministry of Education, Science and Culture (the Ministry of Education, Culture, Sports, Science and Technology since January 6, 2001) under JSPS Fellows (No. 7582).

Appendix

Proof of Theorem 1. The test statistic is a function of the partial sums of $\tilde{x}_t$ as seen in (3). Then using the invariance principle and the continuous mapping theorem we will be able to derive the limiting distribution. Because the invariance principle does not assume a particular distribution, we assume normality of $\{v_t, \varepsilon_t\}'$ in the following. The assumption of normality is convenient to show the limiting distribution as the sum of two independent functionals of Brownian motions, as well as to derive their characteristic functions using the Fredholm approach for Cases 0 and 3.

Firstly, we prove (5) and (7). Because we can easily see that $\tilde{\sigma}^2 \xrightarrow{p} \sigma^2 = \alpha^2\sigma_v^2$ under H_1, where $\xrightarrow{p}$ denotes convergence in probability, we can re-define $S_T = \sigma^{-2}T^{-2}y'MLL'My$ instead of (3) as far as the limiting distribution is concerned.

(i) Let us consider Case 0. Because $[1-DU_t, DU_t]$ spans the same space as $[1, DU_t]$, we can replace $z_t=[1, DU_t]'$ by $z_t=[1-DU_t, DU_t]'$ and then replace the orthogonal projection matrix M by $M_*=diag\{M_a, M_b\}$, where M_a and M_b are the $T_B \times T_B$ and $(T-T_B)\times(T-T_B)$ orthogonal projection matrices on $z_{at}=1$ for $t=1,\dots,T_B$ and $z_{bt}=1$ for $t=T_B+1,\dots,T$, respectively. Then we have the relation $My=M_*y$. Hereafter, we use the subscripts a and b to denote that the vector or the matrix is associated with the data before and after a break, respectively.

The typical jth element of $L'M_*y$ is expressed as $\sum_{t=j}^{T}\tilde{x}_t$, but from the property of the regression we can see that $\sum_{t=T_B+1}^{T}\tilde{x}_t=0$ so that $\sum_{t=j}^{T}\tilde{x}_t=\sum_{t=j}^{T_B}\tilde{x}_t$ for $j \leqslant T_B$. Then we have

$$L'M_*y=L'\tilde{x}=\begin{bmatrix}\sum_{t=1}^{T}\tilde{x}_t\\ \vdots\\ \sum_{t=T_B}^{T}\tilde{x}_t\\ \sum_{t=T_B+1}^{T}\tilde{x}_t\\ \vdots\\ \tilde{x}_T\end{bmatrix}=\begin{bmatrix}\sum_{t=1}^{T_B}\tilde{x}_t\\ \vdots\\ \tilde{x}_{T_B}\\ \sum_{t=T_B+1}^{T}\tilde{x}_t\\ \vdots\\ \tilde{x}_T\end{bmatrix}$$

$$=\begin{bmatrix}L_a' & 0\\ 0 & L_b'\end{bmatrix}\begin{bmatrix}\tilde{x}_a\\ \tilde{x}_b\end{bmatrix}=L_*'M_*y, \tag{A.1}$$

where L_a and L_b are the $T_B\times T_B$ and $(T-T_B)\times(T-T_B)$ matrices with the same structure as L, $L_*=diag\{L_a, L_b\}$, $\tilde{x}_a=[\tilde{x}_1,\dots,\tilde{x}_{T_B}]'$ and $\tilde{x}_b=[\tilde{x}_{T_B+1},\dots,\tilde{x}_T]'$.

Next we decompose the stationary component u_t as $u_t=\alpha v_t+\tilde{v}_{t-1}-\tilde{v}_t$ where $\tilde{v}_t=\sum_{j=0}^{\infty}\tilde{\alpha}_j v_{t-j}$ with $\tilde{\alpha}_j=\sum_{k=j+1}^{\infty}\alpha_k$. Then we can write the stochastic component of y_t as $x_t=\gamma_t+\alpha v_t+\tilde{v}_{t-1}-\tilde{v}_t$, of which the last two terms are asymptotically negligible.

Noting that, in the vectorized form, under H_1, $x=\gamma+\alpha v+\tilde{v}_{-1}-\tilde{v}$ and $\gamma+\alpha v \sim N(0, \sigma_v^2(\alpha^2 I_T+\rho LL'))$, we have, using the relation (A.1),

$$\begin{aligned}S_T &= \frac{1}{\sigma^2T^2}x'M_*L_*L_*'M_*x\\ &= \frac{\sigma_v^2}{\sigma^2T^2}v'(\alpha^2I_T+\rho LL')^{1/2}M_*L_*L_*'M_*(\alpha^2I_T+\rho LL')^{1/2}v+\mathrm{o_p}(1)\end{aligned}$$

$$\begin{aligned}
&\stackrel{\mathrm{d}}{=} \frac{\sigma_v^2}{\sigma^2 T^2} v' L_*' M_* (\alpha^2 I_T + \rho L L') M_* L_* v + o_p(1) \\
&= \frac{1}{T^2} v' L_*' M_* L_* v + \frac{c^2}{\alpha^2 T^4} v' L_*' M_* L L' M_* L_* v + o_p(1), \qquad \text{(A.2)}
\end{aligned}$$

where $v = [v_a', v_b']' = \sigma_v^{-1} (\alpha^2 I_T + \rho L L')^{-1/2} (\gamma + \alpha v) \sim N(0, I_T)$ and $\stackrel{\mathrm{d}}{=}$ denotes equality in distribution. The third relation holds because of the normality of v. From the definition we have

$$L_*' M_* L_* = \begin{bmatrix} L_a' M_a L_a & 0 \\ 0 & L_b' M_b L_b \end{bmatrix}. \qquad \text{(A.3)}$$

In addition, as $M_* L_* v$ is the regression residual of $L_* v$ on the space spanned by $[1 - DU_t, DU_t]$ in the same discussion as Eq. (A.1), the following equivalence holds:

$$L' M_* L_* v = L' \begin{bmatrix} M_a L_a v_a \\ M_b L_b v_b \end{bmatrix} = \begin{bmatrix} L_a' M_a L_a v_a \\ L_b' M_b L_b v_b \end{bmatrix}. \qquad \text{(A.4)}$$

Using (A.2)–(A.4) the LM test statistic is expressed as

$$\begin{aligned}
S_T &\stackrel{\mathrm{d}}{=} \frac{T_B^2}{T^2} \frac{1}{T_B^2} v_a' \left\{ L_a' M_a L_a + \frac{T_B^2}{T^2} \frac{c^2}{\alpha^2 T_B^2} (L_a' M_a L_a)^2 \right\} v_a \\
&\quad + \frac{(T - T_B)^2}{T^2} \frac{1}{(T - T_B)^2} v_b' \\
&\quad \left\{ L_b' M_b L_b + \frac{(T - T_B)^2}{T^2} \frac{c^2}{\alpha^2 (T - T_B)^2} (L_b' M_b L_b)^2 \right\} v_b + o_p(1) \\
&= S_{aT} + S_{bT} + o_p(1), \quad \text{say.} \qquad \text{(A.5)}
\end{aligned}$$

Because v_a and v_b are independent, we can investigate the limiting distributions of S_{aT} and S_{bT} separately.

We first consider the limiting distribution of S_{aT}. Because M_a is the orthogonal projection matrix on z_{at}, we have

$$\begin{aligned}
&\frac{1}{T_B^2} v_a' L_a' M_a L_a v_a \\
&= \frac{1}{T_B^2} \sum_{t=1}^{T_B} \eta_{at}^2 - \left(\frac{1}{T_B^{3/2}} \sum_{t=1}^{T_B} \eta_{at} z_{at}' \right) \left(\frac{1}{T_B} \sum_{t=1}^{T_B} z_{at} z_{at}' \right)^{-1} \left(\frac{1}{T_B^{3/2}} \sum_{t=1}^{T_B} z_{at} \eta_{at} \right), \qquad \text{(A.6)}
\end{aligned}$$

where η_{at} denotes the tth element of $L_a v_a$. By the invariance principle,

$$\frac{1}{\sqrt{T_B}}\eta_{a[T_B r]} = \frac{1}{\sqrt{T_B}}\sum_{j=1}^{[T_B r]} v_j \xrightarrow{\mathrm{d}} B_1(r), \tag{A.7}$$

where $B_1(\cdot)$ is a standard Brownian motion and $[p]$ denotes the largest integer $\leqslant p$. We can also see that

$$\frac{1}{T_B^{3/2}}\sum_{t=1}^{T_B} z_{at}\eta_{at} \xrightarrow{\mathrm{d}} \int_0^1 B_1(r)\,\mathrm{d}r \equiv X(B_1), \qquad \frac{1}{T_B}\sum_{t=1}^{T_B} z_{at}z'_{at} = 1 \equiv \Lambda. \tag{A.8}$$

Then, by (A.6)–(A.8), we have

$$\frac{1}{T_B^2} v'_a L'_a M_a L_a v_a \xrightarrow{\mathrm{d}} \int_0^1 B_1(r)^2\,\mathrm{d}r - X(B_1)'\Lambda^{-1}X(B_1). \tag{A.9}$$

Next, denoting regression residuals of η_{at} on z_{at} as $\tilde{\eta}_{at}$, we have

$$\begin{aligned}
&\frac{1}{T_B^{3/2}}\sum_{t=1}^{[T_B r]}\tilde{\eta}_{at} \\
&\quad = \frac{1}{T_B^{3/2}}\sum_{t=1}^{[T_B r]}\eta_{at} - \left(\frac{1}{T_B}\sum_{t=1}^{[T_B r]} z'_{at}\right)\left(\frac{1}{T_B}\sum_{t=1}^{T_B} z_{at}z'_{at}\right)^{-1}\left(\frac{1}{T_B^{3/2}}\sum_{t=1}^{T_B} z_{at}\eta_{at}\right) \\
&\quad \xrightarrow{\mathrm{d}} \int_0^r B_1(s)\,\mathrm{d}s - Z(r)'\Lambda^{-1}X(B_1),
\end{aligned}$$

where $Z(r)=r$. Because the typical tth element of $L'_a M_a L_a v_a = L'_a\tilde{\eta}_a$ is $\sum_{j=t}^{T_B}\tilde{\eta}_{aj} = -\sum_{j=1}^{t-1}\tilde{\eta}_{aj}$ because $\sum_{j=1}^{T_B}\tilde{\eta}_j = 0$, we have

$$\begin{aligned}
\frac{c^2}{\alpha^2 T_B^4} v'_a(L'_a M_a L_a)^2 v_a &= \frac{c^2}{\alpha^2 T_B^4}\sum_{t=1}^{T_B-1}\left(\sum_{j=1}^{t}\tilde{\eta}_{aj}\right)^2 \\
&\xrightarrow{\mathrm{d}} \frac{c^2}{\alpha^2}\int_0^1\left(\int_0^r B_1(s)\,\mathrm{d}s - Z(r)'\Lambda^{-1}X(B_1)\right)^2\mathrm{d}r.
\end{aligned} \tag{A.10}$$

From (A.9) and (A.10), we obtain

$$\begin{aligned}
S_{aT} &= \frac{T_B^2}{T^2}\frac{1}{T_B^2} v'_a L'_a M_a L_a v_a + \frac{T_B^4}{T^4}\frac{c^2}{\alpha^2 T_B^4} v'_a(L'_a M_a L_a)^2 v_a \\
&\xrightarrow{\mathrm{d}} \omega^2\left\{\int_0^1 B_1(r)^2\,\mathrm{d}r - X(B_1)'\Lambda^{-1}X(B_1)\right.
\end{aligned}$$

$$+\frac{c^2\omega^2}{\alpha^2}\int_0^1\left(\int_0^r B_1(s)\,\mathrm{ds}-Z(r)'\Lambda^{-1}X(B_1)\right)^2\mathrm{dr}\Bigg\}$$

$$=\omega^2 G(B_1;c^2\omega^2/\alpha^2).$$

Precisely in the same way as S_{aT}, we obtain $S_{bT}\xrightarrow{\mathrm{d}}(1-\omega)^2G(B_2;c^2(1-\omega)^2/\alpha^2)$ using $(T-T_B)$ instead of T_B, where $B_2(\cdot)$ is a standard Brownian motion independent of $B_1(\cdot)$, and then (5) is established.

For Case 3, we can replace $z_t=[1,DU_t,t/T,DT_t]$ by $z_t=[1-DU_t,1(t\leqslant T_B)\times t/T_B,DU_t,1(t>T_B)\times(t-T_B)/(T-T_B)]$ and then replace the orthogonal projection matrix M by $M_*=diag\{M_a,M_b\}$, where M_a and M_b are the orthogonal projection matrices on $z_{at}=[1,t/T_B]'$ for $t=1,\ldots,T_B$ and $z_{bt}=[1,(t-T_B)/(T-T_B)]'$ for $t=T_B+1,\ldots,T$, respectively. Then, analogously to Case 0, the relation (5) can be established.

(ii) For Cases 1 and 2, we cannot decompose M as $diag\{M_a,M_b\}$. But in the same way as (A.2) we have

$$\begin{aligned}S_T&=\frac{\sigma_v^2}{\sigma^2T^2}v'(\alpha^2I_T+\rho LL')^{1/2}MLL'M(\alpha^2I_T+\rho LL')^{1/2}v+\mathrm{o_p}(1)\\&\stackrel{\mathrm{d}}{=}\frac{\sigma_v^2}{\sigma^2T^2}v'L'M(\alpha^2I_T+\rho LL')MLv+\mathrm{o_p}(1)\\&=\frac{1}{T^2}v'L'MLv+\frac{c^2}{\alpha^2T^4}v'(L'ML)^2v+\mathrm{o_p}(1).\end{aligned}\tag{A.11}$$

We also have, as (A.9) and (A.10),

$$\frac{1}{T^2}v'L'MLv\xrightarrow{\mathrm{d}}\int_0^1 B(r)^2\,\mathrm{dr}-X(B)'\Lambda^{-1}X(B)$$

and

$$\begin{aligned}\frac{c^2}{\alpha^2T^4}v'(L'ML)^2v&=\frac{c^2}{\alpha^2T^4}\sum_{t=1}^{T-1}\left(\sum_{j=1}^{t}\tilde{\eta}_j\right)^2\\&\xrightarrow{\mathrm{d}}\frac{c^2}{\alpha^2}\int_0^1\left(\int_0^r B(s)\,\mathrm{ds}-Z(r)'\Lambda^{-1}X(B)\right)^2\mathrm{dr},\end{aligned}$$

using $T^{-1/2}\sum_{t=1}^{[Tr]}v_t\xrightarrow{\mathrm{d}}B(r)$, where $B(\cdot)$ is a standard Brownian motion and $\tilde{\eta}_t$ is constructed as $\tilde{\eta}_{at}$ with the full sample. Then (7) is established.

Next, we derive the characteristic function of the limiting distribution. To this end, we use the following lemma. See Theorem 5.13 of Tanaka (1996) for details.

Lemma A. Suppose that the statistic S_T^* is defined by

$$S_T^* = \frac{1}{T} v' B_T v + \frac{\gamma}{T^2} v' B_T^2 v, \tag{A.12}$$

where $\{v\} \sim$ i.i.d.$(0,1)$ and B_T satisfies

$$\lim_{T\to\infty} \max_{j,k} \left| B_T(j,k) - K\left(\frac{j}{T}, \frac{k}{T}\right) \right| = 0, \tag{A.13}$$

with $K(s,t)(\not\equiv 0)$ a symmetric, continuous and nearly definite function. Then S_T converges in distribution and its limiting characteristic function is given by

$$\lim_{T\to\infty} E(e^{i\theta S_T^*}) = [D(i\theta + \sqrt{-\theta^2 + 2i\gamma\theta}) D(i\theta - \sqrt{-\theta^2 + 2i\gamma\theta})]^{-1/2}, \tag{A.14}$$

where $D(\lambda)$ is the Fredholm determinant of K.

To apply the above lemma to S_T, we have to check the restriction (A.13) and whether S_T can be expressed as (A.12).

(i) For Cases 0 and 3, $\omega^{-2} S_{aT}$ has the same expression as (A.12) with $B_T = T_B^{-1} L_a' M_a L_a$ and $\gamma = c^2\omega^2/\alpha^2$. Moreover, after some algebra, we can see that the (j,k)th element of $T_B^{-1} L_a' M_a L_a$ is expressed as $K(j/T_B, k/T_B) + O(T_B^{-1})$ for all j and k where

$$K(s,t) = \min(s,t) - st, \quad \text{and}$$

$$K(s,t) = \min(s,t) - 4st + 3st(s+t) - 3s^2t^2, \tag{A.15}$$

for Cases 1 and 3, respectively, so that both $K(s,t)$'s satisfy condition (30). Then, by Lemma A, the characteristic function of the limiting distribution of S_{aT} is given by

$$\begin{aligned}
\lim_{T\to\infty} E[e^{i\theta S_{aT}}] &= \lim_{T\to\infty} [e^{i(\omega^2\theta)(\omega^{-2} S_{aT})}] \\
&= [D(i\omega^2\theta + \sqrt{-\omega^4\theta^2 + 2ic^2\omega^4\theta/\alpha^2}) \\
&\quad D(i\omega^2\theta - \sqrt{-\omega^4\theta^2 + 2ic^2\omega^4\theta/\alpha^2})]^{-1/2},
\end{aligned}$$

where the Fredholm determinants of (A.15) are given in Theorem 1, as shown by Theorem 6 of Nabeya and Tanaka (1988) and by Eqs. (5.34) and (9.94) of Tanaka (1996, pp. 139, 369).

The characteristic function corresponding to S_{bT} is obtained similarly, and as S_{aT} and S_{bT} are independent, we have expression (6).

(ii) For Case 1, S_T in (A.11) has the same expression as (A.12) with $B_T = T^{-1} L' M L$ and $\gamma = c^2/\alpha^2$, and we can find the kernel $K(s,t)$ satisfying condition (A.13). However, in this case, it seems tedious to work out the Fredholm determinant of K by directly solving the integral equation, and we have to find it by another method. Note here that the characteristic function of the limiting distribution of S_T^* in (A.12) with $\gamma=0$ is given by $[D(2\mathrm{i}\theta)]^{-1/2}$. Then, if we derive the characteristic function of the null distribution of S_T corresponding to the case when $c=0$, we can obtain the Fredholm determinant $D(\lambda)$ by replacing $2\mathrm{i}\theta$ by λ.

To derive the characteristic function under the null hypothesis, we follow the method used by Perron (1991), and use the expression of the limiting distribution (7) with $c=0$. Denote by μ_B and μ_Y the measures induced by the processes $B(\cdot)$ and $Y(\cdot)$ which is generated by the following stochastic differential equation:

$$\mathrm{d}Y(t)=-bY(t)\,\mathrm{d}t+\mathrm{d}B(t), \qquad Y(0)=B(0)=0.$$

Then the measures μ_B and μ_Y are equivalent and the Radon–Nikodym derivative $\mathrm{d}\mu_B/\mathrm{d}\mu_Y$ evaluated at y is given by

$$\mathrm{d}\mu_B/\mathrm{d}\mu_Y(y)=\exp\left[b\int_0^1 y(t)\,\mathrm{d}y(t)+b^2/2\int_0^1 y(t)^2\,\mathrm{d}t\right].$$

(See, for example, Liptser and Shiryayev (1977) and Theorem 4.1 of Tanaka (1996)). Then we obtain

$$\begin{aligned}\phi(\theta;0)&=E\left[\exp\left\{\theta\int_0^1 B(r)^2\,\mathrm{d}r-\theta X(B)'\Lambda^{-1}X(B)\right\}\right]\\&=E\left[\exp\left\{\theta\int_0^1 Y(r)^2\,\mathrm{d}r-\theta X(Y)'\Lambda^{-1}X(Y)\right.\right.\\&\qquad\left.\left.+\frac{b}{2}(Y(1)^2-1)+\frac{b^2}{2}\int_0^1 Y(t)^2\,\mathrm{d}t\right\}\right]\\&=\mathrm{e}^{-b/2}E\left[\exp\left\{\frac{b}{2}Y(1)^2-\theta X(Y)'\Lambda^{-1}X(Y)\right\}\right]\\&=\mathrm{e}^{-b/2}E\left[\exp\{F'AF\}\right]\\&=(\mathrm{e}^b|I-2\Sigma A|)^{-1/2},\end{aligned}$$

where we put $b^2=-2\theta$, $F=[Y(1),X(Y)']'$, $A=diag\{b/2,-\theta\Lambda^{-1}\}$, and Σ is the variance–covariance matrix of F. The last equality follows from the normality of F. Making use of the computerized algebra MAPLE V, we obtain the characteristic function $\phi(\theta;0)=D(2\mathrm{i}\theta)^{-1/2}$ where $D(\lambda)$ is given in Theorem 1.

The characteristic function for Case 2 can be obtained similarly and we omit the proof. □.

Proof of Theorem 2. As in Eq. (A.5) of the LM test statistic, we have

$$S_T^{ps} \overset{\mathrm{d}}{=} \frac{1}{T^2} v_a^{ps\prime} \left\{ L_a' M_a L_a + \frac{c^2}{\alpha^2 T^2} (L_a' M_a L_a)^2 \right\} v_a^{ps}$$

$$+ \frac{1}{T^2} v_b^{ps\prime} \left\{ L_b' M_b L_b + \frac{c^2}{\alpha^2 T^2} (L_b' M_b L_b)^2 \right\} v_b^{ps} + \mathrm{o_p}(1)$$

$$= S_{aT}^{ps} + S_{bT}^{ps} + \mathrm{o_p}(1), \quad \text{say},$$

where $v_a^{ps} = T/T_B v_a$ and $v_b^{ps} = T/(T - T_B) v_b$. Then,

$$S_{aT}^{ps} = \frac{1}{T_B^2} v_a' L_a M_a L_a v_a + \frac{c^2 T_B^2}{\alpha^2 T^2} \frac{1}{T_B^4} v_a' (L_a' M_a L_a)^2 v_a$$

$$\xrightarrow{\mathrm{d}} \int_0^1 B_1(r)^2 \,\mathrm{d}r - X(B_1)' \Lambda X(B_1) + \frac{c^2 \omega^2}{\alpha^2}$$

$$\int_0^1 \left(\int_0^r B_1(s) \,\mathrm{d}s - Z(r)' \Lambda^{-1} X(B_1) \right)^2 \mathrm{d}r$$

$$= G(B_1; c^2 \omega^2 / \alpha^2).$$

Similarly,

$$S_{bT}^{ps} \xrightarrow{\mathrm{d}} \int_0^1 B_2(r)^2 \,\mathrm{d}r - X(B_2)' \Lambda X(B_2) + \frac{c^2 (1-\omega)^2}{\alpha^2}$$

$$\int_0^1 \left(\int_0^r B_2(s) \,\mathrm{d}s - Z(r)' \Lambda^{-1} X(B_2) \right)^2 \mathrm{d}r$$

$$= G(B_2; c^2 (1-\omega)^2 / \alpha^2)$$

and then the limiting distribution of S_T^{ps} can be derived.

The characteristic function of the limiting distribution is obtained analogously to the LM test and we omit the proof. □

Proof of Corollary 1. Firstly, note that, under the assumption of normality and $\rho = \sigma_\varepsilon^2 / \sigma_v^2 = c^2 / T^2$, $\{x_t\}$ satisfies the condition A6b in Bai (1997), so that $T(\omega^* - \omega) = T_B^* - T_B$ is $\mathrm{O_p}(1)$.

To prove the corollary, it is enough to show that

$$\max_{1 \leqslant j \leqslant T} \left| \frac{1}{T^{1/2}} \sum_{t=1}^{j} \tilde{x}_t - \frac{1}{T^{1/2}} \sum_{t=1}^{j} \tilde{x}_t^* \right| = \mathrm{o_p}(1) \quad \text{and} \quad \tilde{\sigma}^2 - \tilde{\sigma}^{2*} = \mathrm{o_p}(1), \tag{A.16}$$

where $\tilde{x}_t^*$ and $\tilde{\sigma}^{2*}$ are defined in the same way as $\tilde{x}_t$ and $\tilde{\sigma}^2$ with the estimated break point T_B^* instead of T_B. Note that

$$\sum_{t=1}^{j}(\tilde{x}_t - \tilde{x}_t^*) = \sum_{t=1}^{j}\left\{ z_t^{*\prime}\left(\sum_{i=1}^{T} z_i^* z_i^{*\prime}\right)^{-1}\sum_{i=1}^{T} z_i^* z_i^{\prime}\beta - z_t^{\prime}\beta\right\}$$
$$+\sum_{t=1}^{j}\left\{ z_t^{*\prime}\left(\sum_{i=1}^{T} z_i^* z_i^{*\prime}\right)^{-1}\sum_{i=1}^{T} z_i^* x_i\right.$$
$$\left. - z_t^{\prime}\left(\sum_{i=1}^{T} z_i z_i^{\prime}\right)^{-1}\sum_{i=1}^{T} z_i x_i\right\}, \tag{A.17}$$

where z_t^* is defined in the same way as z_t with the break date T_B^*. For the first term, we can easily show that $T^{-1}\sum_{i=1}^{T} z_i^* z_i^{*\prime}$ has the same limit as $T^{-1}\sum_{i=1}^{T} z_i^* z_i^{\prime}$ because $\omega^* = T_B^*/T \xrightarrow{\text{p}} \omega$. In addition, we can see that $\sum_{t=1}^{j}(z_t^{\prime}\beta - z_t^{*\prime}\beta)$ is $\mathrm{O_p}(1)$ for all j. For example, in Case 3, we have

$$\left|\sum_{t=1}^{j}(z_t^{\prime}\beta - z_t^{*\prime}\beta)\right| \leqslant \sum_{t=1}^{T}|(z_t - z_t^*)^{\prime}\beta|$$
$$\leqslant |[0, |T_B - T_B^*|, 0, |T_B - T_B^*|]\beta| = \mathrm{O_p}(1),$$

because $T_B - T_B^* = \mathrm{O_p}(1)$. This shows that the first term of (A.17) is $\mathrm{O_p}(1)$ for all j.

The second term of (A.17) can be decomposed into

$$\sum_{t=1}^{j} z_t^{\prime}\left(\sum_{i=1}^{T} z_i z_i^{\prime}\right)^{-1}\left(\sum_{i=1}^{T} z_i^* x_i - \sum_{i=1}^{T} z_i x_i\right)$$
$$+\sum_{t=1}^{j} z_t^{\prime}\left\{\left(\sum_{i=1}^{T} z_i^* z_i^{*\prime}\right)^{-1} - \left(\sum_{i=1}^{T} z_i z_i^{\prime}\right)^{-1}\right\}\sum_{i=1}^{T} z_i^* x_i$$
$$+\sum_{t=1}^{j}(z_t^* - z_t)^{\prime}\left(\sum_{i=1}^{T} z_i^* z_i^{*\prime}\right)^{-1}\sum_{i=1}^{T} z_i^* x_i. \tag{A.18}$$

Note that $\sum_{i=1}^{T} z_i^* x_i - \sum_{i=1}^{T} z_i x_i$ is $\mathrm{o_p}(T^{1/2})$ while $\sum_{i=1}^{T} z_i z_i^{\prime}$ and $\sum_{t=1}^{j} z_t$ are $\mathrm{O_p}(T)$ for all j. Then, the first term of (A.18) is $\mathrm{o_p}(T^{1/2})$ for all j. Similarly, the second and third terms can also be shown to be $\mathrm{o_p}(T^{1/2})$ for all j. As a result, the second term of (A.17) is $\mathrm{o_p}(T^{1/2})$ for all j and then $\max_j |T^{-1/2}\sum_{t=1}^{j}(\tilde{x}_t - \tilde{x}_t^*)| = \mathrm{o_p}(1)$. Thus the first relation of (33) is established.

In the same way, the second relation of (A.16) can be proved. □

References

Andrews, D.W.K., 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. Econometrica 59, 817–858.

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821–856.

Andrews, D.W.K., Ploberger, W., 1994. Optimal tests when a nuisance parameter is present only under the alternative. Econometrica 62, 1383–1414.

Bai, J., 1994. Least squares estimation of a shift in linear processes. Journal of Time Series Analysis 15, 453–472.

Bai, J., 1997. Estimation of a change point in multiple regression models. Review of Economics and Statistics 79, 551–563.

Banerjee, A., Lumsdaine, R.L., Stock, J.H., 1992. Recursive and sequential tests of the unit-root and trend-break hypothesis: theory and international evidence. Journal of Business and Economic Statistics 10, 271–287.

King, M.L., Hillier, G.H., 1985. Locally best invariant tests of the error covariance matrix of the linear regression model. Journal of the Royal Statistical Society (B) 47, 98–102.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., 1992a. Testing for stationarity in the components representation of a time series. Econometric Theory 8, 586–591.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992b. Testing the null hypothesis of stationarity against the alternative of a unit root. Journal of Econometrics 54, 159–178.

Leybourne, S.J., McCabe, B.P.M., 1994. A consistent test for a unit root. Journal of Business and Economic Statistics 12, 157–166.

Liptser, R.S., Shiryayev, A.N., 1977. Statistics of random processes I: general theory. Springer, New York.

Nabeya, S., Tanaka, K., 1988. Asymptotic theory of a test for the constancy of regression coefficients against the random walk alternative. Annals of Statistics 16, 218–235.

Nelson, C.R., Plosser, C.I., 1982. Trends and random walks in macroeconomic time series. Journal of Monetary Economics 10, 139–162.

Nunes, L.C., Kuan, C.M., Newbold, P., 1995. Spurious break. Econometric Theory 11, 736–749.

Park, J.Y., Sung, J., 1994. Testing for unit roots in models with structural change. Econometric Theory 10, 917–936.

Perron, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57, 1361–1401.

Perron, P., 1990. Testing for a unit root in a time series with a changing mean. Journal of Business and Economic Statistics 8, 153–162.

Perron, P., 1991. A continuous time approximation to the unstable first-order autoregressive process: the case without an intercept. Econometrica 59, 211–236.

Perron, P., 1997. Further evidence on breaking trend functions in macroeconomic variables. Journal of Econometrics 80, 355–385.

Perron, P., Vogelsang, T.J., 1992a. Nonstationarity and level shifts with an application to purchasing power parity. Journal of Business and Economic Statistics 10, 301–320.

Perron, P., Vogelsang, T.J., 1992b. Testing for a unit root in a time series with a changing mean: corrections and extensions. Journal of Business and Economic Statistics 10, 467–470.

Tanaka, K., 1996. Time Series Analysis: Nonstationary and Noninvertible Distribution Theory. Wiley, New York.

Vogelsang, T.J., 1997. Wald-type tests for detecting breaks in the trend function of a dynamic time series. Econometric Theory 13, 818–849.

Vogelsang, T.J., Perron, P., 1998. Additional tests for a unit root allowing for a break in the trend function at an unknown time. International Economic Review 39, 1073–1100.

Zivot, E., Andrews, W.K., 1992. Further evidence on the great crash, the oil-price shock, and the unit root hypothesis. Journal of Business and Economic Statistics 10, 251–270.